

14 October 1982

What policy for West Germany?

The new government in Bonn was promising last week to keep faith with research and development, which is honourable and even creditable. But it should not forget the problems hitherto ignored.

The new ministers of the federal government of West Germany are still forgivably startled to find themselves where they are. Even though the disintegration of ex-Chancellor Helmut Schmidt's coalition government had been apparent since the Hamburg elections at the end of last year, its final collapse last week could not be accurately predicted. So it is inevitable that the new ministers should begin their spells of office by saying what their party has been saying in opposition. This explains what Dr Heinz Riesenhuber, the new minister of research and technology, was saying in Bonn last week (see page 569). Yes, the role of the large national research establishments must be re-examined, and links with industry must be strengthened. (Technology transfer is the catchword.) Yes, industry itself must shoulder more of the cost of research and of responsibility for its direction, but the federal government will help with tax incentives and by other means. And more of the cost of keeping students at universities will be financed by loans rather than by grants. All this is consistent with what the Christian Democrats have been saying in opposition, but it will be some time before the new government can break new ground, with policies that go beyond the generalities of opposition. Indeed, that time may be postponed by the new Chancellor's promise of general elections in March next year, politically contentious and potentially distracting.

Luckily, West Germany has one important circumstance in its favour: the new government party echoes what the beaten Social Democrats have always said, that strong government support for basic research is essential to future prosperity. The trouble, so far, is that the funds spent on research have been used inefficiently (see page 570). The crucial question now is whether the new government will have the political courage and skill first to acknowledge that reforms are necessary and then to carry them out. People in West Germany, in the research community and elsewhere, have been agonizing for the past several years about the reasons for what appears to be poor productivity in research. All kinds of fanciful explanations have been offered, from the aftermath of the Third Reich to the partition of Germany thirty years ago. The real trouble is too much organization (see *Nature*, 23 May).

One obvious impediment to change is that the ministry in Bonn which has been for the past decade the most powerful influence on science and technology in West Germany has no direct influence on what happens in universities, where the need for change is most urgent (and, on recent form, the ministry has been almost indifferent to the problems of academic research). Another and more serious difficulty is that the federal government has no direct influence on the universities, which are constitutionally the creatures of the *Länder* governments in whose regions they happen to be situated. For the new government to require, as it should, that the traditional autonomy of the select company of tenured professors (*Ordinarien*) should be moderated by, say, peer review, assessment of the use of research funds, or that student numbers should be restricted at the more popular universities, would be to take on a series of dog-fights, especially with the *Länder* governments still in the hands of the Social Democrats. Sooner rather than later, however, these battles must be fought.

Fortunately, the federal government has some levers more immediately in its hands. While the constitutional difficulties would be formidable, a first priority should be an attempt to

engineer a redistribution of university research spending from university subventions by the *Länder* governments, now all short of funds, to central grant-making organizations such as the Deutsche Forschungsgemeinschaft. The result would be a concentration of research in universities and departments with proven track records. Another urgent need is somehow to bridge the long-standing gulf between technically excellent Max Planck institutes and the rest of the academic research system. For while there continue to coexist two constitutionally separate, even if occasionally geographically contiguous, sets of institutes for basic research, the system is bound to be less effective than it could be. Even where Max Planck institutes work closely in collaboration with university researchers (which is not always the case), the fact that the people concerned follow distinct career tracks is a continuing cause of resentment and a source of needless rigidity. Will the new ministers of education and of research have the wit to see what could be accomplished by joint action?

The new government has a more untidy problem in deciding what to do about the major national laboratories that have emerged since the end of the Second World War. The rhetoric of the past few days, like that of the past thirteen years, is that the national laboratories must be judged by the contribution they make to industrial innovation and development. In reality, these federal laboratories are too mixed a bunch for that test to be applicable. Thus the cancer research centre at Heidelberg may in passing be a source of industrial innovations (in the techniques of body scanning, for example), but this is only accidental. Similarly, the heavy-ion establishment at Darmstadt (the laboratory that synthesized an isotope of element 109 a few weeks ago) should not be judged by its immediate contribution to industrial change but as a source of basic research as such (and thus comparable with the accelerator laboratory at Hamburg). What the federal government needs, therefore, is not a uniform policy for persuading these establishments to be more directly of benefit to West German industry but a clear understanding that they have different functions. Some of them could with advantage be managed by the Deutsche Forschungsgemeinschaft rather than the Ministry of Research and Technology, for then their role as parts of the academic research enterprise would be more apparent and more easily developed. Others, the nuclear research establishments at Jülich and Karlsruhe for example, may need more drastic treatment.

The temptation in the next few months will be for the new government to put these claims on its attention to one side, seeking instead ways of implementing its election promises about industrial tax incentives or resisting (as is its political inclination) demands for special help for the electronics industry or some other suffering from overseas competition. The snag is that it is rarely possible in the real world to compensate for the effects of one set of structural defects by devising a different set of institutional regulations. Incentives for industrial research are not to be scorned, in West Germany, although there will be many who would prefer to wait for evidence on the working of the Reagan Administration's schemes, introduced just over a year ago. But it is even more important that the new government should face up to the problems, administrative, legal and even constitutional, that have everybody crying about the disappointments of the past few decades while pretending that the causes are unknown.



London's agony (contd.)

The University of London may surprisingly have become a model of good administration.

The University of London is the British university system in microcosm, but with the added disadvantages that its budget is more vulnerable to the flight of high fee paying students from overseas and that its administration is poorly suited to making decisions, even of the simplest kind. This week, the university has told its constituent parts how much money from the centre they will have to spend in the financial year that began more than two months ago, on 1 August. On this occasion, however, the delay is partially forgivable, for the court of the university, formally responsible for money matters that do not involve questions of academic policy (whatever they may be), has been uncustomarily decisive. In particular, it has broken out of the mould in which it had become trapped of sharing its funds as they had happened to be shared in the previous year, with allowances for special pleading on the side. Even so, much remains to be done.

The university's constitutional difficulty stems from the notional autonomy of the parts of which it is composed. Formally, the university cannot tell one of its constituent colleges how to conduct its affairs except by insisting through a cumbersome network of committees that acceptable academic standards should be followed and by doling out funds in such a way as to encourage or discourage various activities. (In dealings from the centre with individual universities in the United Kingdom, the University Grants Committee responsible for sharing out the government subvention of universities can only command the second weapon, at least in the short term.) Two years ago and then last year, the university somewhat arbitrarily told its colleges what they would have to spend in the succeeding years without explaining how the figures had been decided, except for the still puzzling and unwarrantable decision that courses combining two different lines of study should be discouraged. This year, the university has told its dependants how it has done its sums.

To its credit, the university has also devised an important innovation, and has decided that constituent colleges able to attract substantial amounts of research income should be paid extra on that account. This decision is entirely consistent with all that has been said in recent months about the collapse of what is called the dual-support system, by means of which the general university subvention is meant to cover the overhead cost of carrying out research (but nobody likes the word "overhead"). This is the most tangible way in which the university can substantiate its frequently expressed hope that research will survive all the troubles there have been. The failure to follow such procedures in recent years has unfairly robbed of funds those colleges with strong traditions in research but which for other reasons have been dealt with meanly within the university.

It is also sensible that in this first year, the university should have followed the logic of its conclusion that research requires extra support only part of the way, to the tune of one-third of what will be done two years from now. One of the ironies of the redistribution in the coming year is that the only one of the nine regular undergraduate schools to gain funds as a result, Queen Elizabeth College, is one of those dealt with most severely last year. (Imperial College deals separately with the University Grants Committee, while the London School of Economics has no substantial amount of research in scientific subjects.) The end result of this new arrangement is certain to be that the different parts of the university will have an incentive to improve their research. Should not the British system as a whole now promptly follow suit?

The university has also been sensible in planning for a reduction of the number of its parts. While it remains to be clearly demonstrated that small colleges with only about a thousand or so students are inviable, there is something in the view that concentration would help and might be cheaper. That course of action is at least worth a try. So colleges with plans to merge with

each other are now offered a modest incentive — an increase of combined student numbers — and possibly a share of the pot of money the university is salting away meanwhile. Exactly such incentives would have allowed the University Grants Committee, in its dealing with the British universities collectively, to help rationalize the system without as much pain as there has been. In London, with the university's new-found taste for open government, what remains to be done is to explain exactly why those colleges that will find life even more difficult if they choose to stand alone have been singled out for such a fate, and to make more durable arrangements than now exist for making sure that potential casualties of the rearrangement, people and institutions, are not discarded without thought.

Soviet science not quoted

Why are Soviet papers less often cited by US authors than in the 1970s? Secrecy? Or quality?

Is Soviet science approaching a crisis of quality? Or is the regime becoming more and more secretive about its good science? These questions, which have an important bearing on international science policy and questions of technology transfer, are attracting much attention. Although the data to answer them are scant, some trends are arresting. Thus in the field of mathematics, US scientists have been citing Soviet work less and less; US citations of Soviet mathematics declined by 48 per cent from 1973 to 1979 (see *Nature* 30 September, p.388). More recent unpublished data from the National Science Foundation suggest that the decline is now 53 per cent.

Why would US authors make less and less use of Soviet mathematics — a field in which that country is believed to excel? The principal translators of Soviet mathematical work in the United States are the American Mathematical Society, in Providence, Rhode Island, and Plenum Press in New York. (The latter is a principal source of translating and distributing foreign scientific materials, especially materials from Eastern bloc countries.) Both organizations report that, if anything, they are translating and distributing more Soviet mathematics papers than they did a decade ago. So the decline in citations is not apparently connected with a decline in the amount of material available in translation in the United States. Another explanation has unexpectedly turned up in a study for the US Department of State carried out by Computer Horizons Inc., a New Jersey corporation. Francis Narin, president of the corporation, says that there are two distinctly different groups of highly cited Soviet scientific papers: those describing work of the highest quality which is cited over and over again, and those which are cited often by Soviet scientists but much less often elsewhere. It seems as if Soviet authors look to one group of Soviet mathematicians while Western authors look to another. The study was too preliminary to show whether the work in the two sets of papers is related, but Narin has established that the papers cited by Westerners were not merely translations of those cited most by the Soviet scientists. The results so far are, however, a telling illustration of what citation analysis may eventually accomplish in this important field.

Meanwhile, speculation about the causes of these changes flourishes. To many US academics, it seems natural to account for what is happening as a consequence of Soviet classification policy. The case of computer science and its applications is often mentioned: in the United States, most work not strictly related to defence and commercial development is published freely, while in the Soviet Union the opposite is true. But the past decade has also seen more modest declines in the frequency of US citations of Soviet work in physics and the Earth and space sciences, while US citations of work in these fields by other than Soviet nationals have risen dramatically. So there is again speculation that the Soviet Union may be hiding its best work. For the time being, US academics do not accept that there has been a decline of quality of Soviet research. But they are intrigued at the possibility that citation analysis could help them find out.

Threat to French research directors

Ructions about twelve-year tenure limit

Senior scientists in France were staggered last week to learn that they may remain as research directors — heads of laboratories or research groups — for no more than 12 years. The new regulation, announced by science and industry minister Jean-Pierre Chevènement, will be applied retroactively from 31 December 1985. Hence any director who has now been in post for more than nine years must contemplate resignation by that date.

There have been strong reactions, both positive and negative. Professor Jean Dausset, the Nobel prize-winning immunologist, has described it as “a good measure”, mobility being necessary at all levels of research. “Our laboratories are dying of sclerosis”, he said. But according to another world-renowned biologist, the measure is “very bad for French science”.

In fact, the regulation is not yet law, and may be modified. François Gros, ex-director of the Institut Pasteur and now science adviser to the Prime Minister, Pierre Mauroy, is being urged from all quarters to advise Chevènement against the move.

The chance that such pressure will succeed, however, seems slim. Chevènement has in fact already *hardened* the proposal from the version which originated from the medical research council, INSERM. INSERM probably suffers more than most French research councils from the longevity of its research directors, some of whom have been in the job for between 20 and 25 years. It was generally agreed that the worst should be shaken out, to leave room for new blood. The question was how?

After long debate stimulated by the new director-general of INSERM (Philippe Lazar, who was *rapporteur* for the National Colloquium on Science and Technology last January), it was agreed by the directorate, the scientific council and the trade unions at INSERM that the 250 or so posts of scientific director should be limited to four-year terms, renewable twice (for 12 years in all). After that, all agreed that the director of a group could be re-confirmed in his post “in exceptional circumstances”. But when the ministry was consulted about the proposal, it insisted that the 12-year term would be final, and that there would be no exceptions.

According to one leading scientist close to the INSERM directorate, the result is a harsh, bureaucratic solution to a real problem in France: how to make effective evaluations of scientific progress. “It’s

true I’ve never seen a group leader fired”, he said. But the imposition of a rigid 12-year rule is not the answer. Rather, there should be proper, expert evaluation, using both French scientists and scientists from abroad. Many existing INSERM directors should go, he felt; but equally others who had been in place 12 years should stay for the good of science such as the neuro-physiologist Jacques Glowinski, or Michel Jouvet (a specialist in sleep mechanisms), both threatened by the new rule.

Unfortunately for French science the whole issue is rapidly becoming politicized, with those resisting the move being classed as rightist. However, left-wing scientists also consider the move hypocritical — because it avoids the “democratic” assessment systems set up by the new government. These systems, which involve largely elected committees to make judgements on appointments, grants and so on, are sup-

posed to be “transparent” and avoid the formation of cliques and interest groups. But if they are expected to work, why does the government not make use of them to decide whether a group leader should continue or not? The imposition of a rigid rule reduces the power of the committees, and indicates that the government lacks confidence in their decisions.

Whatever final decision is reached on the 12-year rule, INSERM is at pains to smoothe the highly troubled waters. On Monday it released a statement arguing that very few of its directors would be affected (a medical newspaper in Paris last week estimated that 60–70 positions were under threat); and that the object of the reform was above all “the quality of research”. The period between now and the deadline of 31 December 1985 would be one of “intense reflection on how to apply the rule”, the statement said.

Robert Walgate

Few changes for West Germany

Heidelberg

The new West German Minister of Research and Technology, Dr Heinz Riesenhuber, gave a good impression last week, after 48 hours in office, of having hit the ground running. Once party spokesman on energy, he was obviously familiar with his role, fielding even questions on the fast breeder with confidence and humour.

Riesenhuber holds a PhD (University of Frankfurt, 1965) for work on lattice distortions in microcrystalline iron phosphate. For eleven years, he has been the technical director of a thriving middle-sized company specializing in industrial polymers. He entered the Bundestag in 1976 and became chairman of the intra-parliamentary committee on energy and environment in 1977.

Last week he stressed the deteriorating national economic situation, the increasing financial burden of the prototype reactors (see *Nature* 30 September, p.384) and the inertia of large projects which limit the immediate possibility of increasing investment and change. But Riesenhuber sees a positive role for his ministry in the implementation of party research and technology policies (*Nature* 23 September, p.289).

Nevertheless, Riesenhuber emphasized his commitment to basic research and to continuing support along the lines of the previous administration. But why, he wondered, is world-class investment in basic research not always matched by research results, as assessed by criteria such as the *Science Citation Index* or Nobel prizes?

The new minister has already carried out an analysis of potential growth areas. Besides microelectronics and biotechnology, materials research, particu-

larly high-temperature polymers, was mentioned. (There are suggestions of a new Max Planck Institute for polymer research.)

To promote these key areas, Riesenhuber will encourage closer collaborations with other ministries, research organizations and industry, so as to create “market push” rather than “government pull”. He plans to take a close look at the country’s many large research organizations which, despite their solid research achievement, may not be pulling their weight economically.

Riesenhuber also stressed the importance of integrating risk assessment — financial and environmental — at all stages of scientific and technological undertakings. The *post factum* imposition of controls resulted in expense, delays and environmental damage that increased public disquiet about the role of science and technology in a vulnerable world. This, he says, is the only way to restore public trust in science and technology. How far he succeeds in dispelling the environmentalists’ fears, the polls will tell.

By contrast with Riesenhuber’s bustle, all was peaceful last week at the sister Ministry of Education and Science where Dorothee Wilms has taken over as minister. But the new administration’s decision to put student grants on a loan basis will have far-reaching effects. There are no signs of French expansionism, but, although a general cut of 5 per cent in all expenditure is expected, Gerhard Stoltenburg, the new Finance Minister and himself once a forceful minister of scientific research under Kiesinger in the 1969 Christian Democrat government, may yet exempt science and technology and even provide the 5 per cent increase needed to

cover inflation. Playing for safety, however, Riesenhuber plans also to encourage industry to finance more research. He considers that large projects can be better initiated and carried out by industry than by the cameral system of federal government. It remains to be seen how well he succeeds — and what the consequences would be for the government's own research establishments. **Sarah Tooze**

Why the lag?

Why has research in West Germany been a disappointment? Economically, West Germany has been if anything better off than its competitors. But West Germany has problems peculiar to itself, as follows:

- Length of education. Both school and university education take longer in West Germany than elsewhere, so that the average age for beginning a PhD is 27. The results are expense and loss of some of the most creative years.

- The PhD is finished at an age when tenured jobs and stability are more attractive than mobility, so grants for travel are undersubscribed.

- Too few untenured posts. Under German employment laws there must be a reason for ending a job. Despite loopholes and contrivances, mobility within the research community is limited.

- Legal responsibility of the *Ordinarius* (C₄) professors. By law, responsibility for finance and administration is usually vested only in the C₄ professor. This creates a sharply pyramidal hierarchy which deprives other members of staff of responsibility.

- Research budgets of universities are contributed three quarters by the region (*Land*). This money goes automatically to the C₄ professors, who have complete control and who are not subject to review of any kind. Federal money allotted by a normal grant system covers one quarter of university research spending and is the only money allocated subject to review.

- The 1976 university law laid down that all grant money, even when allocated to a specific person for a specific purpose, must be administered by the university. This has spawned a bureaucratic machine which leaves the grant-user trapped in a mesh of regulations and subject to interference. These procedures discourage industry from financing research in the universities.

- Inefficient administration. There is a lack of professionalism in the university administrations and administrators are rarely appointed within university departments to help research staffs and coordinate with the university.

- The numbers of students have soared, but despite overall increased spending, research money per head has fallen.

Sarah Tooze

Polish universities

More signs of latent discontent

"Universities must not turn into an arena for political games", the Polish government newspaper *Rzeczpospolita* warned as the new term opened on 1 October. Government and party officials were evidently worried that the reassembly of some 400,000 young people in the country's 87 universities and higher colleges could provide a focus for further confrontations — particularly in the atmosphere of heightened tension following the enactment of the new law on trade unions which dissolves all existing union structures (including the 11-million strong Solidarity). In the event, the students, whose own Independent Students' Association (NZS) was outlawed last January, reassembled peacefully, and the problem now facing the universities is not so much the threat of student political action but rather of inaction.

Calling for a "climate of normality" in the universities, the party daily *Trybuna Ludu* stated last week that "it will be difficult for the academic milieu to fulfil its task if it does not embark on a dialogue on such issues as the new law on higher education, improvement in upbringing and teaching methods, and the future of the student movement".

The President of the Council of State, Henryk Jablonski, reiterated this theme when conferring professorial diplomas in Warsaw on 6 October. Normal functioning of society will be impossible, he stressed, if citizens are "inactive", and students should therefore be trained to show "initiative, independent thinking and imagination". A conference of university lecturers from Lower Silesia and Opole went so far as to urge encouraging a "creative" approach by students to political and social studies (courses once more obligatory for all).

The students, however, seem reluctant to embark on dialogue with an academic establishment they no longer trust. Some 30 university rectors, elected under the previous democratic procedures have been replaced under martial law, and, under the political "verification" procedures, some 2 per cent of university staff have failed to meet the required criteria, while more than 5 per cent have been re-employed only conditionally, and 2½ per cent have had to take up non-teaching jobs within the universities.

"Internal emigration" — passive non-cooperation with the political and social life of the country — became widespread among Polish intellectual and creative circles under martial law, and has been widely advocated in underground Solidarity literature. In what appears to be a last-ditch stand to "channel" the aspirations of the students, the authorities are now trying to establish a new students' organization, to replace both the banned

NZS and the party-linked Socialist Union of Polish Students (SZSP), a move that many SZSP activists themselves urge as a way to bring the mass of students back into the party fold.

Trade union law, which effectively put an end to hopes for a return to the liberalizations of 1980–81 has already moved some Polish students to consider turning their internal emigration into an external one, and to try to continue their studies abroad. Some Western embassies in Warsaw, the British Embassy in particular, have already foreseen this possibility and are granting visas to students only if financial support is assured by a Western sponsor (see p.574). Such sponsors will not prove easy to find. In the United Kingdom, where a Polish Students' Appeal Fund was set up to help students stranded by the declaration of martial law, universities have been generous with free places, but funds for maintenance and such necessities as textbooks are, in the words of Dr Stanislaw Gomulka of the London School of Economics, "already over-extended".

Vera Rich

Canada's research plans

Holding up

Washington

Canada's ambitious five-year plan for a rapid growth of investment in research and development is in trouble — but it is surviving present economic difficulties more successfully than almost any other government initiative.

This mixed picture emerges from a report just issued by the Canadian National Research Council (NRC), the government's national laboratory. NRC's budget for the 1981–82 fiscal year fell short of the \$304 million it said it needed to carry out even the most restrictive of three alter-

OECD data on research and development expenditure

Country	Year	Expenditure as % of GDP
United States	1980	2.49
Switzerland	1979	2.45
Germany	1979	2.27
United Kingdom	1979	2.20
Japan	1980	2.04
Netherlands	1978	1.97
Sweden	1977	1.90
France	1979	1.79
Belgium	1979	1.40
Norway	1979	1.37
Yugoslavia	1977	1.20
Canada	1981	1.18*
Finland	1979	1.08
Denmark	1979	0.97
Italy	1979	0.82
Turkey	1978	0.59
Portugal	1978	0.32
Greece	1980	0.18

*The OECD figures include social sciences and humanities research. The figures cited in the text (1.07 per cent current and 1.5 per cent goal) do not.

natives in the 1980 five-year plan. That plan is aimed ultimately at doubling the production of technology-intensive industry by 1990.

Even so, the \$296 million that NRC spent in 1981–82 was a 30 per cent increase over the previous year, allowing NRC to go ahead with several new programmes, including the Industrial Materials Research Institute in Boucherville, Quebec, an ice research project in Newfoundland and a small tokamak fusion research programme in cooperation with Hydro-Quebec.

NRC officials are worried, however, that the government will not be able to

sustain the annual 30 per cent increases called for in the plan. Since budget items can be approved by the government throughout the fiscal year, the 1982–83 budget will not be known until the year is over, but officials are predicting that it will fall significantly short of the planned figure. The big worry is that what NRC calls “core support” is going to suffer the most. This is the reserve of expertise and backbone research facilities (analytical chemistry laboratories, for example) that NRC says must be built up if the government is to be able to address problems as they arise.

NRC officials say the tendency has been for immediate needs — such as the recent “fire-fighting” operation on the health-effects of urea-formaldehyde foam insulation — to draw personnel and resources away from the core. This, they say, “jeopardizes the viability of NRC as Canada’s national laboratory”. It seems likely also to jeopardize the nation’s effort to boost its overall research and development by 1985 from 1.07 per cent of gross national (GDP) to 1.5 per cent — and thus to bring it closer in line with the more industrialized countries of Western Europe.

Stephen Budiansky

Bilingual planners disagree

Brussels

Two separate plans to create an advanced electronics industry in Belgium have just been announced, one by the new Flemish regional government and the other by the central government’s Minister for Science, Philip Maystadt. Although there are claims that the two plans complement one another, it seems likely that Belgian research and development policy will again become a source of conflict between the country’s two communities.

Like the French plan and the European Commission’s *Esprit* programme, the new initiatives spring from fears that European industry is in danger of missing the boat for the “Third Industrial Revolution” as the Flemish are calling it.

Since last December, Belgium’s already complex political arena has been further complicated by the delegation of much of central government’s power to two separate regional governments, one for Flanders and the other for Wallonia. The micro-electronics strategy presented by the president of the Flemish “government”, Gaston Geens, is one of the first fruits of this constitutional change.

It will cost some BF 4,000 million (£44 million) over the next four years. Unlike the French plan, it is strongly orientated towards stimulating private industry and differs in its emphasis on the production of customized chips. A large part of the funds will go towards establishing with Bell Telephone (a Belgian affiliate of ITT) a silicon foundry manufacturing customized microchips. The Belgian company is putting up half the cost of BF 2,800 million (£32.5 million) and the Flemish government the remainder.

The second element will be the major expansion of the research facilities at Leuven University, led by Professor Overstraeten, into a laboratory for advanced research into micro-electronics at a cost of some BF 1,500 million. This will be complemented by a programme of grants to engineers already working in industry or finishing doctorates to improve their knowledge of the state of the art of very large-scale integrated chips and other fields. To help industry directly, there will be a consultancy centre where companies

can receive advice on how to computerize either their products or their production.

The Flemish micro-electronics programme has strong backing from the region’s political parties, and industry has organized itself into pressure groups such as FLORA — Flemish companies interested in the development of the information technology industry.

The central government plan, still to be officially adopted, concentrates on creating a receptive environment for companies in the forefront of electronics. Above all, Maystadt proposes to use the public purchasing power of the state in the defence and telecommunications sectors to favour Belgian companies, particularly those seeking their first customer. This is

not strictly permitted under EEC or GATT rules, but Belgium awards more public works contracts to foreign companies than any other EEC country.

BF 200 million will also be set aside for subsidizing research and development work in companies involved in developing customized chips, in the belief that Belgium will never be able to compete with the French or Americans in producing standardized chips. A further BF 240 million will be spent on a study on the sociological impact of micro-electronics in an effort to calm the Walloon trade unions hostile to the changes their use will bring.

The central government’s programme is estimated to cost BF 820 million a year but this figure may be revised in the light of efforts to cut the large government deficit.

Jasper Becker

US medical profession

Fight against anti-trust laws

Washington

The first of October may turn out to have been a decisive date for the medical profession in the United States, ending a shoot-out between the profession and the Federal Trade Commission (FTC), the government agency that seeks to monitor anti-competitive, anti-consumer practices in the medical profession (as well as others).

The issue, according to FTC’s chairman, James C. Miller III, is whether the professions should be immune from the normal inquisitiveness of the agency empowered to protect consumers. “Admission to a profession should be a guarantee of competence, not a guarantee of immunity from the laws the rest of us must obey”, Miller has said.

As long ago as December 1975, FTC issued a complaint against the American Medical Association (AMA), the profession’s powerful umbrella organization, barring it from prohibiting advertising by doctors, among other things. Since then, AMA has fought the complaint in the courts. But when the US Supreme Court was deadlocked on the issue earlier this year, the FTC order automatically became effective on 1 October, the first day of fiscal year 1983.

But AMA has not given up. Having failed in the courts, it seeks relief in Congress by amending FTC’s authorizing legislation to prevent it from proceeding against the professions. The changes stood a good chance of being passed — reflecting AMA’s considerable power in Washington — until Congress went into recess. AMA spokesmen still say they can change the law when Congress returns in November. But Congress’s huge agenda, its lame-duck status and the Reagan Administration’s support for FTC in this matter may combine to frustrate AMA’s plans.

The fight has been long because several interlocking issues are involved: whether professionals can be the subject of anti-trust action, whether FTC has the power to regulate the professions, whether AMA has reformed itself sufficiently in the past seven years to be considered now above reproach. To complicate matters, FTC has issued related orders against the American Dental Association (which agreed to be bound by the outcome of the AMA battle), the ophthalmologists, and several state and local medical societies. All these support AMA in its fight with FTC.

But the central issue is who should determine medical practice locally in the United States. FTC’s position is that state and

local medical societies have enormous influence over doctors and over state legislatures and courts. Thus they can facilitate doctors' boycotts of hospitals that charge too little, prevent physicians from working on a salary instead of the more lucrative fee-for-service basis and decide whether to encourage or to discourage advertising.

Regulation of these professional groups is traditionally the job of the states. But, FTC argues, it is hard for state authorities to rule against these powerful groups. Many consumer groups believe that the influence of these professional associations is connected with their powerful political fund-raising ability. In May, a bill introduced by Thomas A. Luken (Democrat, Ohio) that would place a moratorium on all FTC action against the professions had 155 co-sponsors in the House of Representatives. According to Congress Watch, a public interest group, those 155 co-sponsors had received \$599,000 in campaign contributions from AMA, \$191,000

from the American Dental Association and \$41,000 from the ophthalmologists. By October, the bill had 221 co-sponsors.

A more limited measure was attached to FTC's authorizing legislation last May by the Senate Committee on Commerce, Science and Transportation. It would declare that FTC has no jurisdiction over state-regulated professionals and their professional associations. This measure may well remain in the final Senate bill.

William Rial, president of AMA, is unsparing in his criticism of FTC's allegedly blundering intrusion into the medical profession which, he writes, is "threatening the future of medical practice in this country". The drain on medical societies' resources because of FTC administrative procedures, Rial says, "can hinder these groups from performing their most important functions. . . . The standards of quality established by the American Medical Association and other medical societies . . . are being undermined by a federal

agency that possesses no medical qualifications." AMA spokesmen point out that FTC has investigated some professional groups only to find nothing at all wrong after a great expense of time and money. An AMA spokesman in Washington says the profession is already feeling the chill of the FTC order. Local cost-containment groups, he says, are afraid to meet "for fear FTC will investigate them for price-fixing".

But this seems to be just what FTC wants. It cites the case of fees for eye glasses and optometry, where the agency won the right to promote advertising. In a short time, the price of eye glasses and contact lenses fell by as much as a half. The cost of health care in the United States continues to rise in double figures, even though inflation now stands at approximately 6 per cent. FTC is not promising that it can lower the overall costs but it now has a licence to try to do so.

Deborah Shapley

Commonwealth agriculture

More journals mean more food

The Commonwealth Agricultural Bureaux (CAB) are among the first beneficiaries of the British government's new International Organizations' Act passed early last year. Incorporation under the new act has given CAB full international status and thus removed some of the previous obstacles to legal agreements between the bureaux and national governments. Hence a new station for the Commonwealth Institute of Biological Control in Kenya can now go ahead because the Kenyan government has a legal entity with which to sign a contract. The new act also means that CAB member governments, which contribute to the financial upkeep of the bureaux, need no longer fear that part of their contribution will end up in the British exchequer. CAB employees, most of whom work in Britain, no longer pay income tax.

After nearly 30 years gestation, CAB was finally born in 1930, since when it has grown into a major international information service for agricultural scientists. It offers three basic services — identification of biological specimens, advice on biological control and the publication of agricultural abstracts and some primary journals. Three Commonwealth institutes — those of entomology, helminthology and the mycological institute — provide identification services. The Commonwealth Institute of Biological Control, the only CAB institution based outside Britain, provides advice on biological control from its headquarters in Jamaica. The four institutes and a further ten bureaux prepare abstracts of the literature in their own specialities and conduct literature searches for enquirers.

Until 1975, CAB had been largely financed by contributions from member

governments. Since then, however, the publishing activity has been made self-financing by substantially increasing the charge for journals. The biological control service also earns half its keep from contracts. But the identification and taxonomy services are financed from member governments' contributions.

Of the £6 million needed to run CAB each year, £4.5 million is now earned, mainly from journal sales and charges for literature searches. (CAB's abstracts data-

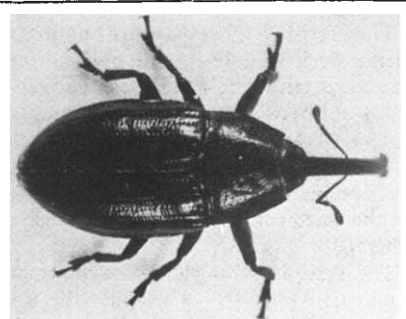
base is now available on-line on the world's major host computers.) The remaining £1.5 million is provided by member governments which contribute roughly according to their means. Thus, Britain makes the largest contribution of about £400,000 (or just 30 per cent of £1.5 million), Canada, Australia and New Zealand come next and many of the smaller Commonwealth nations contribute only a little.

Member nations are entitled to a quota of free journals depending on the size of their contributions, and to a slightly discounted rate for extra journals, on-line services and advice on biological control. Non-member nations may also take advantage of CAB's services — but they must pay slightly over the odds. The identification service, however, is free for all, members and non-members alike.

The United States and Japan, two non-member nations, have in fact made greater use of CAB's services than any other nation. Thus, CAB are open to the criticism that many developing countries are subsidizing services which the developed countries are best equipped to use. That criticism was reinforced when CAB made their abstracts available on-line on the world's major host computers, all of which are in the industrialized north.

At a conference in 1980 to review CAB's work and celebrate their golden jubilee, developing countries expressed a wish for greater access to CAB's services but also their appreciation of them. For although the developed countries may benefit most — Britain as the host country most of all — CAB often provide developing countries with their only access to the world agricultural literature or to advice on agricultural pests. The alternative for developing countries — to finance their own identification and biological control services and establish their own libraries — would in most cases be prohibitively expensive.

Judy Redfearn



This weevil, *Elaeidobius kamerunicus*, should save Malaysian agriculture around \$115 million a year if things go according to plan. The weevil (actual length 4 mm) has been introduced into Malaysia from West Africa and has already greatly increased the fruit production of Malaysia's oil palms and eliminated the need for hand pollination of the trees. Since its introduction in 1981 the weevil has spread throughout the Malaysian peninsula and increases of 20 per cent in yield have been obtained. As a spin-off from this work, based on a study by Dr R.A. Syed of CAB's Commonwealth Institute of Biological Control, growers in West Africa are now keen to find means of controlling oil palm pests biologically — since the continued use of insecticides will inevitably limit the effectiveness of the pollinating weevils.

Nobel prizes 1982

Prostaglandins

The 1982 Nobel Prize for Medicine was awarded this week jointly to Dr John R. Vane of the United Kingdom and Dr Sune Bergström and Dr Bengt Samuelsson of Sweden for their "discoveries concerning prostaglandins and related biologically active substances".

It has taken fifty years for prostaglandins to be deemed worthy of a Nobel prize. They were discovered in semen by the Swede Ulf van Euler who misnamed them prostaglandins believing them to originate from the prostate whereas, in fact, they came from the seminal vesicles. Sune Bergström's interest in prostaglandins stemmed from the early contributions of van Euler who must be counted unlucky not to share in the prize. Bergström, at the Karolinska Institute in Stockholm used newly developed chromatographic methods for the separation of lipids for the first purification of prostaglandins. Bergström and J. Sjövall published the empirical formulae of prostaglandins E and F in *Acta chemica scandinavica* in 1960. Two years later, Bergström and Bengt Samuelsson, produced the first complete chemical structure of a prostaglandin.

The structure suggested that arachadonic acid, an essential fatty acid, was a precursor of prostaglandins. Bergström and Samuelsson proved that point in 1964 but only with the help of David van Dorp of Unilever who provided them with radioactively labelled arachadonic acid; van Dorp published the same result independently but simultaneously. Samuelsson has continued to make important contributions, most notably in determining the structure of the prostaglandin-related thromboxanes and leukotrienes.

Two major contributions are attached to John Vane's name. The first is the discovery, published in *Nature* in 1971, that the anti-inflammatory effects of aspirin and other non-steroidal drugs are related to their ability to inhibit the synthesis of prostaglandins. Vane was then at the Institute of Basic Medical Sciences but later went to the Wellcome Research Laboratories where he discovered prostacyclin, a highly unstable prostaglandin which can inhibit platelet aggregation. The work on prostacyclin was primarily in the hands of Salvador Moncada, first author of the key paper (*Nature* 263, 663; 1976).

Pharmacologically, prostaglandins have marked effects on the contraction of smooth muscle and they are implicated in uterine and blood vessel contraction. They have been used in the induction of labour and abortion in humans and are used widely in veterinary practice to synchronize oestrus prior to artificial insemination of a herd of cows.

Peter Newmark

Biotechnology index

Cool reception for Genex shares

Washington

The Genex Corporation of Rockville, Maryland, one of the larger US biotechnology companies, went public on schedule by offering 2 million shares of stock for public sale on Wednesday 29 September (too late for inclusion in *Nature's* September Biotechnology Index, see table).

As always, the first public stock offering is a major event for a young company as a test of how sound and promising the business community believes it to be. And as the first public offering of a major biotechnology corporation for some time, Genex's debut was closely watched by those concerned with genetic engineering as an investment. Wall Street analysts say that the fact that Genex could soon derive steady income from the artificial sweetener market, and that it has work under way in the interferon field and other promising areas, made the stock trade well. But the lack of a market "hype" or a well-publicized human drug product made the Genex offering lacklustre, in the opinion of some analysts.

Even so, the stock traded quickly at \$9.50 per share and all 2 million shares were sold. Thus, Genex grossed approximately \$19 million, and after fees can now put around \$17 million in the bank.

Genex was founded by J. Leslie Glick, a former senior cancer research scientist at Roswell Park Memorial Institute. Glick had previously, in 1969, founded Associated Biomed Systems, Inc., which filed for bankruptcy in 1977. In 1977 Glick started Genex with backing from Koppers Company Inc. of Pittsburgh and the InnoVenn Group of venture capitalists. At the time of the offering, Koppers held 45 per cent of the company's shares. InnoVenn

held 22 per cent, Robert F. Johnston, chairman of the board, held 11 per cent and Glick had 11 per cent. Now, Glick's 1,080,000 shares in Genex have a book value of approximately \$9 million.

The offering also gave the company's founders some inkling of what their shares are worth. Genex has grown steadily and now has 211 employees. Each year, it has increased its income through research contracts. In mid-1982, Genex had eight major research contracts: four with Japanese companies (including work on interferon for Green Cross) and two with Swedish companies. A seventh contract was with Koppers and the eighth was with Bristol-Meyers. The company has a pilot plant producing L-aspartic acid for possible use in producing aspartame (an artificial sweetener considered to have great market potential — see *Nature* 16 September p.199).

The market's calm response to the offer of Genex shares showed that the US business community no longer views genetic engineering as the next best thing to the Midas touch. Genex earlier planned to sell 2.5 million shares for up to \$12 per share, and could have grossed as much as \$33 million. But it was felt that the market would not bear such a price, so the offer was pitched lower and the company's plans scaled back. After a week of trading, Genex stock was selling at about \$8.50 per share. By contrast, in 1980, Genentech, another stable, well-regarded firm, offered its first public stock at \$35 per share and the price rose, in hours of frenetic trading, to \$89 per share. By late September, however, Genentech stock was selling below the original offer price, reflecting the market's calmer mood.

Deborah Shapley

Nature index of biotechnology stocks

1982 high	1982 low	Company	Close previous month	Close 24 Sept	Change
34 1/4	16 1/8	A.B. Fortia (Sweden)	32 3/4	34 1/4*	+ 1 1/2
8	2	Bio Logicals (Canada)	2	2 1/2	+ 1/2
7	3 3/8	Bio-Response (USA)	4	4 3/4	+ 3/4
14 1/8	8	Cetus (USA)	8 7/8	8 3/4	- 1/8
11	6 1/8	Collaborative Research (USA)	8 3/8	8 5/8	+ 1/4
21 7/8	14	Collagen (USA)	17 5/8	18 1/2	+ 7/8
8 7/8	5 3/4	Damon (USA)	7 7/8	7 1/2	- 1/4
17 1/4	11 1/4	Enzo-Biochem (USA)	12	14 1/2	+ 2 1/2
28	6 5/8	Flow General (USA)	10 5/8	10 7/8	+ 1/4
37 3/4	26	Genentech (USA)	33 3/4	33 1/4	- 1/2
3 3/8	2 1/4	Genetic Systems (USA)	3 1/8	3 1/4	+ 1/8
17 7/8	9 3/8	Hybritech (USA)	14 3/8	13 1/2	- 7/8
10 3/4	6 1/4	Molecular Genetics (USA)	10 3/4	10 3/4*	0
46 3/4	34 7/8	Novo Industri A/S (Den.)	46 3/4	44	- 2 3/4
12 1/8	8	Monoclonal Antibodies (USA)	9 3/4	8 1/4	- 1 1/2

The *Nature* Biotechnology Index for September 1982 stands at 117.9, compared with 118.1 last month. Base is 100 as of 25 June 1982. Previous indexes appeared *Nature* 12 August p.599 and 9 September p.101. Close-of-month prices are the closing prices on the last Friday of each month. Where stocks are traded over the counter, the price quoted is the bid price. For stocks traded on the American and New York Stock Exchanges, the price quoted is the transaction price. Data from E.F. Hutton, Inc.

*High or low for this calendar year.

CORRESPONDENCE

SERC and South Africa

SIR — The replies to the questionnaire sent out by a working group of the Science and Engineering Research Council (SERC) on the South African Astronomical Observatory (SAAO) should not be taken as representing the balance of opinion among astronomers on whether the United Kingdom should continue to subsidize the SAAO (see *Nature* 23 September, p.291). The questionnaire was not sent to all UK astronomers and was in fact sent mainly to groups who have already used the SAAO. Many of those who refuse to use SAAO on principle did not receive it. The questions asked were all concerned with what facilities should be used, what programmes pursued and what instrumentation required. As there was no question on whether the subsidy to SAAO should continue, I and many others did not reply to the questionnaire, regarding it as an irrelevance to the central issue.

The central issue is a moral one, whether astronomy is so important that it should come before one's duty to the oppressed black majority in South Africa, whose leaders (those who have not been murdered or imprisoned) have repeatedly called for the severing of all international ties with South Africa. This call has been supported by the United Nations, UNESCO and many other international bodies. The agreement maintained by SERC with the South African government on behalf of UK astronomers is probably the last official cultural link between the two countries and is a disgrace to UK science.

Note that I am not trying to prevent individuals with strong stomachs from using telescopes in South Africa or for that matter Chile. The issue is whether a substantial chunk of the UK astronomy budget should go to maintain an official subsidy of SAAO by the United Kingdom.

I ask all concerned scientists, and UK astronomers particularly, to write to the Chairman of the Astronomy, Space and Radio Board, c/o SERC, North Star Avenue, Swindon, calling for an end to this shameful treaty.

MICHAEL ROWAN-ROBINSON

Department of Applied Mathematics,
Queen Mary College,
London E1, UK

How to help Poland

SIR — As a recent and relatively frequent visitor to Poland, I have the unhappy task of reporting to the Western scientific community the particularly unfortunate situation that Polish scientists now find themselves in.

Specifically, they are caught in a pincer between the scientific policies of the Polish and US governments. On the Polish side, in this time of economic crisis during which the average Polish standard of living has already decreased by 25 per cent, it is a fact that the relative decrease in salaries for scientists is far greater than for most other economic groups. For example, under the new system coal miners' earnings are about four times that of professors and more importantly, coal miners have access to stores that have something for

sale other than the most basic foods while professors do not.

The reasons for this especially poor treatment of scientists by the politically captive Polish government are: (1) universities and institutes have generally been organizational centres for Solidarity; (2) the scientific community is generally Western-oriented in its values and has been particularly tainted by extensive contacts and visits to the West; (3) this Western orientation precludes its usefulness to the Eastern bloc defence establishment; (4) the Soviet Union has little intrinsic interest in the cultural value of Polish science.

Salaries, prices and availability of goods have been restructured to the great detriment of scientists. Polish scientists, thanks to money saved during visits to the West, have lived relatively well in the sense of better apartments, clothes and cars. Their Polish salary has only paid for subsistence at a moderate level. The reduced salaries now barely cover monthly expenses in normal markets, where little is available, and are certainly very inadequate on the black market where such luxuries as decent soap, toothpaste and shirts are to be purchased. Clearly the future is economically even more downhill for Polish scientists than it is for the general worker. The second arm of the pincer is the American policy of cutting support for joint US-Polish projects. This money was very important in that it financed travel to Western scientific conferences. The Polish government rarely releases its foreign exchange for such purposes. The US policy is based on the idea that actions which punish Polish science and culture will cause Polish scientific and cultural

groups to apply pressure on the Polish government, thereby affecting its policies.

Sensible as this may seem in countries whose governments serve their people, as we see from the above discussion the policy cannot be effective in Poland. In fact the policy is damaging to Western interests in that it helps destroy one of the most Western-oriented communities in Poland.

Can we in the West do anything to help our Polish colleagues? The answer is yes. This correspondent has already communicated the contents of this letter to the US government at the highest levels. On the economic and intellectual side, invitations to established, as well as young, Polish scientists to come to the West, even for the more easily arranged short visits, can be a great help. It will keep the Poles abreast of Western scientific work. It will also be amazingly effective in helping them survive economically. To appreciate this one must realize that \$100 saved when converted to zloty on the ever present and unofficially tolerated black market (the government wants the dollars) is roughly equal to four to five months salary for a scientist at the beginning of his scientific career. Importantly, it has been my experience that Polish visitors return excellent value for their salaries. They truly value their chance to work in the luxury of Western scientific facilities and because of the limited time they have, they work very hard.

Finally, invitations to visit the West help elevate the spirit of our Polish colleagues who truly live these days under the heavy foot of external oppression.

Name and address
supplied but withheld

No rising trend for IQ in Japan

SIR — R. Lynn¹ reported evidence of IQ levels in Japan and drew some far reaching implications for American and European economies. In his letter, T.B.L. Kirkwood (*Nature* 2 September)² suggested, rather cautiously I thought, that the evidence presented does not entirely support a "rising trend".

Taking the data presented for the mean IQs of the 23 Japanese cohorts, I plotted the cumulative sum chart³ shown in the figure. Without making any statistical test it is clear that there are two "populations" and no

obvious outliers. There is certainly no evidence for a "rising trend" in any part of this data.

As to likely explanations, in the absence of further information regarding the way that the data were gathered, I suggest that the following hypotheses can be advanced:

- (1) There was step change in the ability of the Japanese to do intelligence tests starting with those born after 1946.
- (2) There was a step change in the test applied to those born after 1946.

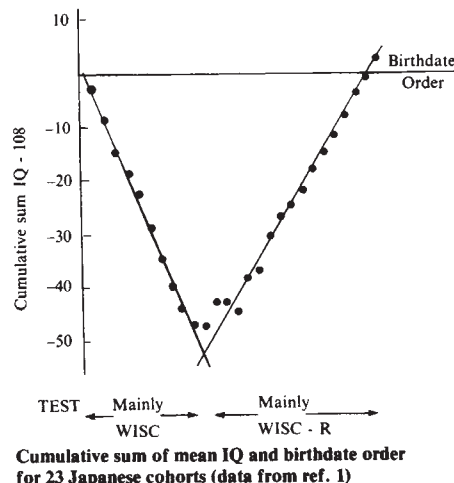
Before getting too involved in the implications of hypothesis (1) it seems prudent to investigate hypothesis (2) more thoroughly, for while it remains open, doubt must be cast on IQ comparisons such as those made by Lynn in his paper, and indeed on the levels themselves.

Whether the Japanese are more intelligent than other nations or whether IQ measures anything that can be related to "economic success", are no doubt fascinating questions. What the Japanese have done is to allow and encourage a broad mass of people to use their intelligence in their ordinary jobs, and in particular by many national training programmes since 1950, to deal with data effectively by means of simple charts and statistical treatments.

E.J. PEARSON

Manchester, UK

1. Lynn R., *Nature* 297, 222-223 (1982).
2. Kirkwood, T. B.L. *Nature* 299, 8 (1982).
3. *Cumulative Sum Techniques* (ICI Monograph No. 3, Oliver and Boyd, Edinburgh, 1964).



NEWS AND VIEWS

Limitations of Maxwell's distribution

A calculation of the evolution of a system of gravitating masses (galaxies) in an expanding universe surprisingly suggests that Maxwell's velocity distribution is not an acceptable end point.

EVERYBODY knows that if you put a collection of massive interacting particles in a box and let them settle down, in due course their velocities will have a maxwellian distribution. The interaction may be only sporadic, physical collision, as in a classical monatomic gas for example, or otherwise be ever-present, as with the long-range forces of gravitation. Either way, the end result is that the chance that the velocity of a particle lies between v and $v + dv$ is proportional to

$$v^2 \exp(-\frac{1}{2}mv^2/kT) dv$$

where m is the particle mass, T the temperature and k is Boltzmann's constant. For what it is worth, the proportionality constant is $4\pi(m/2\pi kT)^{3/2}$. Maxwell's first flawed derivation of this distribution function in 1860 was afterwards recast along the lines of Ehrenfest's much later principle of detailed balancing — at equilibrium, the rate at which particles are removed from some velocity interval will be equal to the rate at which particles with such velocities are produced, by collision or otherwise. By 1878, Boltzmann had put kinetic theory on an analytical basis with his integro-differential equation admitting only permissible distribution functions as solutions. During the First World War, Enskog and then Chapman showed how the maxwellian distribution is the end point in the evolution in time of an arbitrary initial velocity distribution, at least if interactions between the particles dissipate no kinetic energy. So what happens if the interacting massive particles are not atoms of a gas but a collection of galaxies (for simplicity, with equal mass) and if the box into which they are put is the expanding universe?

Certainly not a maxwellian distribution, as can be told from the treatment of this problem by Henry E. Kandrup of the University of California, Santa Barbara, published in an August issue of *Astrophysical Journal* (259, 1; 1982). The problem is of course relevant to the current preoccupation in cosmology with the formation of clusters of galaxies. While many hold to the theory of Zel'dovich and his colleagues that galaxy clusters are formed by the internal coalescence of large pancake-shaped aggregations of matter, recent Monte Carlo simulations of the evolution of gravitating masses suggest that clustering may be a natural outcome of a different starting point — a collection of galaxy-sized objects in an expanding box. Kandrup's study is an attempt at an analytical treatment of the problem hitherto dealt with primarily by numerical modelling. The most striking of his conclusions is that if you put a collection of identical massive objects (galaxies) in an expanding box (universe) and take proper account of their long-range interaction through gravitation, then you find that the total kinetic energy of the massive objects increases monotonically with time. Naturally, the steady increase of kinetic energy must be compensated for somehow — by what Kandrup calls the "generation of negative correlational energy", presumably another name for potential energy. In short, the calculation suggests that clustering should be a natural occurrence, while the steady increase of kinetic energy is a direct proof that a maxwellian distribution is entirely out of court.

On a little reflection, this conclusion should not be as surprising as it seems. For the expanding universe is hardly a system in equilibrium, and since Enskog's inaugural dissertation at the

University of Uppsala in 1917, it has been in principle possible to calculate deviations from simple maxwellian distributions in the many familiar physical circumstances in which equilibrium does not obtain, the transport of heat through a kinetic gas for example, or the transport of momentum as in viscous phenomena. Kandrup acknowledges the way in which Prigogine and his associates have more recently dealt with non-equilibrium systems, showing how circumstances can arise in which a system can be deflected from relaxation towards a maxwellian distribution. With all this said, however, the physical implications of these phenomena are not at all well understood. Is the time ripe for revival of interest in classical kinetic theory?

Kandrup's analysis of his own problem of what happens to a collection of gravitating point masses in an expanding universe contains a further and perhaps deeper surprise. The calculation starts from a time-dependent distribution function for the values of each of the independent variables (space coordinates and velocities) for each of the galactic masses, whereupon it is possible to relate the time-rate of change of the distribution function to the space gradient of the distribution function and to the forces between particles. The procedure is strictly analogous to the direction of Liouville's theorem, and says little more than that probability is conserved. The algebra, however, is complicated by the need to work in an appropriate and moving reference frame. Exactly by analogy with classical kinetic theory, Kandrup nevertheless sets out to calculate a one-particle distribution function by integrating over the coordinates of all other masses and arrives at the equivalent of Boltzmann's equation for the time-rate of change of that same function. Analytical solutions of this equation are for the time being out of reach, if only because of the complications flowing from the $1/r$ form of the gravitational potential, but there are approximations that put the analogue of Boltzmann's equation into barely manageable form. And Kandrup's conclusion is that in these circumstances, a maxwellian distribution of velocities is not a stationary solution of the equation. And, Kandrup insists, this unexpected result "has nothing to do with the fact that the universe is expanding" but is directly a consequence of the long-range character of the gravitational forces between all the particles.

The implications of Kandrup's calculation for the understanding of the clustering of galaxies remain to be seen. The fact that the monotonic increase of kinetic energy which tumbles out of the calculations implies a decrease of potential energy which is at least consistent with the formation of galaxy clusters is of at least more than passing interest. But Kandrup's approximations deserve to be scrutinized with great care by those with an interest in his problem of galaxies in an expanding box as a test of simple classical kinetic theory. Hitherto, attempts to calculate velocity distribution functions in circumstances where particles interact by means of inverse-square law forces — in ionized gases, for example — have tended to be frustrated by the recognition that 'collisions' are not simple binary collisions when such long-range forces are involved. If Kandrup has found a way round this kind of difficulty, he will no doubt be embraced not merely by cosmologists but by plasma physicists as well. □

An immunoregulatory molecular complex with five active sites

from N.A. Mitchison

THE cellular immunology group at Yale University has just published a very important claim¹. It comes in several parts, of which the most important is to have discovered a new suppressive transmitter or 'factor' with five distinct active sites.

Antibody production in the immune response is believed to be regulated in part by a cascade of interacting T cells which produce soluble transmitters that act on others in the chain and ultimately suppress the activity of the helper T cells that induce the proliferation of the antibody-producing B cells. The new transmitter is made by two cells at the end of the suppressor T cell cascade, each of which releases a single peptide chain (or, just possibly, a very stable assembly of chains). The two chains assemble to make up a complex which acts in a suppressive manner on the target of suppression, a helper T cell. Two of the five active sites on this two-chain molecule are necessary for assembly of the chains. One of these sites is a product of the H-2J locus of the major histocompatibility complex, made by a cell which bears the same product on its surface. H-2J is well known as a marker on suppressor T cells and their products, so it is no surprise to find it turning up here. The other site is on the other chain made by an H-2J⁺ cell, and it binds the H-2J product on the H-2J⁺ chain.

The H-2J⁺ chain also has an anti-idiotypic site, which binds to an idiotype on the surface of the target cell. Idiotypes are antigenic determinants of the antigen-binding sites of antibodies. They can also be detected on T cells where they are believed to be associated with the antigen-binding site of the T cell receptor. Because of their association with the antigen-binding site, they are thought to be the target of antigen-specific regulation of the immune response. However, it is a little surprising to encounter them at the end of the suppressor cascade, since they have more usually been associated with its induction rather than with its final action. (Gershon *et al.* do not use the term idiotype; I suspect this is because they believe that products similar but not identical to immunoglobulin are involved, which are encoded by a family of related genes on the same chromosome and immediately to the right of the immunoglobulin genes. Whether or not they are correct in this belief, the idiotypic nomenclature seems appropriate and is in general use for markers identified by the mouse strains they use.)

The other chain, in addition to the H-2J-

binding site, carries an antigen-binding site and an effector site that mediates the suppression of the target cell. The antigen-binding site presumably makes a bridge via antigen to the target cell, and thus helps the transmitter find its appropriate target which it then suppresses, Gershon *et al.* suggest, by cleaving to generate a toxic peptide. They have in mind the way diphtheria toxin acts, and more immediately a cleavage mechanism to yield a toxic product which has recently been demonstrated for the products of a T suppressor cell clone².

It is further proposed that this transmitter is only one among a class of similar molecular assemblies, all of which are composed of a pair of chains, one with an anti-idiotypic binding site and the other with effector activity. In another publication³, the Yale group describes another such factor which acts earlier in the suppressor cascade. The function of this one is to induce activity in the terminal cells of the cascade, and accordingly its effector site has an activating rather than a suppressive effect. Thus, as a general rule, the anti-idiotypic chain can be regarded as a 'schlepper' or carrier, targetting the effector activity to the T cell bearing the appropriate antigenic specificity.

The importance and originality of this part of the claim is as follows. In the first place, it provides fresh insight into the linkage between cells within the immunoregulatory circuits and cascades. If transmitters really work through generating active peptides in the way that Cantor, Gershon and their colleagues propose, then a whole new field of cell biology opens up concerning the mode of action of these peptides. It will be reminiscent of the study of bacterial toxins and colicines, but different insofar as it will deal with physiological interactions rather than attack and defence. Furthermore, the idiotype- and antigen-binding activity of these transmitters makes sense only if they act at a fair distance. It is refreshing to find this kind of evidence at a time when the case for interleukin-2, hitherto the strongest candidate for a long-distance transmitter between T cells, is weakening^{4,5}. What is more generally agreed is that it profits little to debate whether these transmitters operate physiologically or are *in vitro* artefacts. After all, the neurophysiologists have been debating the same question over substance P for the last fifty years, and the molecular biologists seem to be on the point of building enough of these immunologically active transmitters to answer some really decisive questions. This, in fact, is another reason why the present claim is so important: the new idiotype- and H-2J-binding sites should

provide useful handles for grabbing these polypeptides in transcription experiments.

But surely what is most important is to have brought together the idiotope- and antigen-recognizing systems. Hitherto authorities such as Sercarz and Janeway (himself a member of the Yale group) have argued that these two activities are mediated by different T cells ('Ts1 and Ts2'), acting independently of one another and converging only at the end of the regulatory cascade. What is now proposed is that within the suppressor cascade at least, the two activities are not separate.

This brings us to the second part of the claim, advanced in the present publication with further details promised later, that the anti-idiotypic repertoire of T cells is acquired in the thymus. Thus the presence of an (A × B) F₁ thymus graft in an A-strain mouse enables its T cells to produce a factor which recognizes immunoglobulin-linked products on B-strain cells under circumstances where normal A-strain T cells would not be able to do so. This again runs contrary to current doctrine, that the T cell receptor repertoire for recognition of idiotypes is largely independent of the thymus.

How convincing are these claims? The idea that a single factor may carry both J- and an antigen-binding site was implicit in the very first work on J, and has been extended recently to include a factor with J-binding and antigen-binding activities⁶. Incidentally, if Zembala *et al.* are correct, the target of the J- and antigen-binding factor is itself still a suppressor cell, not a helper. What is new is that five activities can all go together, including two from distinct chromosomes (H-2J is on chromosome 19, immunoglobulin on chromosome 12) that appear to be on the same chain. The Yale group indeed recognizes this difficulty, and has no solution to offer at present. They cite other work on two-chain suppressor factors, but the analogy is surely misleading since these involve chains separately encoded by chromosomes 12 and 19.

Perhaps the main point is this. There is a further part of this study which is still owing. In the publications which I have so far seen, there is no direct stoichiometry of the proposed molecular assembly. Obviously the two chains need to be characterized separately, and their binding to one another and to appropriate substrates examined, either in solution or on appropriate absorbents. It looks as though some of the data could be acquired rather easily, particularly as significant readings can be made from as little as "three individual calculations from each culture condition". Perhaps the Yale group already has these data. Until they are available, the group's claims should be treated with some caution. □

1. Flood, Yamauchi & Gershon *J. exp. Med.* (in the press).
2. Fresno *et al.* *J. exp. Med.* 153, 1260 (1981).
3. Yamauchi *et al.* *J. exp. Med.* (in the press).
4. Keene & Forman *J. exp. Med.* 155, 768 (1982).
5. Crispe *et al.* *Transplant Proc.* (in the press).
6. Zembala, Asherson & Golizzi *Nature* 297, 411 (1982).

N.A. Mitchison is in the Department of Zoology, University College London, London WC1E 6BT.

Mapping the Local supercluster

from Joseph Silk

WHAT can be learned about the mechanism of galaxy formation from observations of the Local supercluster, the array of some thousands of galaxies, including those of the Virgo cluster, which is the only one easily accessible to astronomers? The question has been made pointed by the now popular view that only large-density fluctuations, containing 10^{14} solar masses and which are therefore perhaps 10^3 times as massive as the Milky Way system, can have survived from the early universe. Studies of the structure of the Local supercluster (confirming, among other things, that there is indeed a structure) have a bearing on this theory, suggesting that the supercluster may be the evolved state of a minimal surviving density fluctuation.

The prevailing view is that in the early universe, small-scale structure was suppressed by the diffusive smoothing due to radiation to which matter was coupled or, if massive neutrinos were dominant, by collisionless phase mixing and damping. Only when matter could move freely and when it first combined into atoms, some 3×10^5 years after the big bang, was gravitational instability effective at enhancing the growth of density.

Pressure gradients would have been negligible at this stage, and Zel'dovich and his colleagues in Moscow have argued that matter in a region of density fluctuation whose magnitude was of order unity over its extent would probably collapse to what would initially be a highly anisotropic pancake-like configuration. Recent numerical simulations suggest that both filamentary and sheet-like structures may be the generic outcome of such collapses, with great voids separating the highly compressed regions to form an almost cellular structure. The characteristic scale is about 5 Mpc if baryonic matter predominates, but may be somewhat larger if massive neutrinos play that role.

On this theory, the dense sheets of compressed gas cool rapidly by radiation and become unstable to further fragmentation. Galaxies form from the fragments, initially in flattened or elongated configurations which are likely to be only transient, because continuing gravitational collapse and infall tends to make the galaxy distribution spherical. However if the universe is open, with a present mean density of about one-tenth the critical value for closure, gravitational forces do not significantly modify the pancakes if the initial collapse is fairly recent, say at a redshift less than 10.

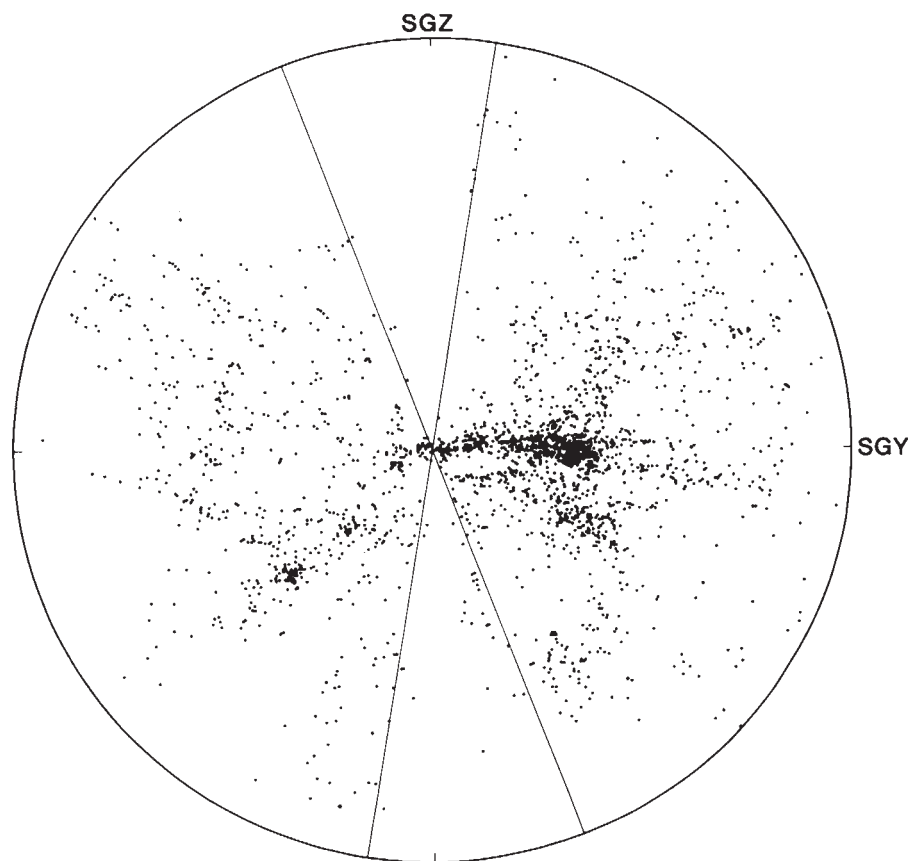
To what extent is this theory consistent with what is known of the distribution of galaxies? Recently, data bearing on the

theory have come from three-dimensional maps of the galaxy distribution and from redshift surveys. There has been a substantial advance on earlier studies of galaxy counts, which yielded only two-dimensional information. Vast regions devoid of luminous galaxies have been discovered, and the galaxies are often distributed in great chains of clusters and superclusters. This is certainly in qualitative agreement with the pancake theory, but it remains difficult to make quantitative comparisons of predictions with observations.

New studies of the galaxy distribution in our immediate neighbourhood are accordingly timely, and may provide the closest and most detailed view of large-scale structure. It has been known for many years that there is a clear excess of nearby galaxies (above background) in a region that extends over 90° or more near the north galactic pole and called by De Vaucouleurs the Local supercluster. The Local Group of galaxies, of which our own Milky Way galaxy and the Andromeda galaxy Messier 31 are the most prominent, lies on the outskirts of the Local supercluster.

The most detailed chart so far of the Local supercluster has been compiled by R. Brent Tully (*Astrophys. J.* **257**, 389; 1982), who measured some 2,200 galactic redshifts extending to about three times the distance of the Virgo cluster. He finds that the Local supercluster possesses structure and comprises two distinct components — a thin disk containing sixty per cent of the luminous galaxies with the remainder in an inhomogeneous halo. The thickness of the disk is only 2 Mpc (if the Hubble constant is $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$), whereas its diameter is 12 Mpc. In the halo component, almost all luminous galaxies appear to be in a small number of diffuse clouds, so that most of the halo volume is empty. The more prominent clouds of galaxies are elongated, and their long axes are directed towards the centrally situated Virgo cluster.

To understand this remarkable structure, Tully argues that the disk of the Local supercluster can be so thin only if random motions of galaxies normal to the disk are less than 100 km s^{-1} or if there are very substantial amounts of unseen matter dominating the gravitational attraction within it. The prolate structures found in the halo are suggestive of tidal interactions soon after the halo clouds were formed. The disk thinness is very suggestive of the pancake theory, as is the existence of large volumes devoid of galaxies, but the halo



Inferred positions of 2,175 galaxies projected on a plane roughly coincident with the plane of the Milky Way system. The wedge-shaped regions are the exclusion zones determined by the Milky Way. The dense patch of galaxies to the right of centre includes the Virgo cluster. The radius of the circle, inversely proportional to Hubble's constant, is 30 Megaparsec if Hubble's constant is 100 km s^{-1} . (From Tully *Astrophys. J.* **257**, 389; 1982.)

Joseph Silk is Professor of Astronomy at the University of California, Berkeley, California 94720.

clouds pose more of a problem. The mass of a cloud is only about 10^{13} solar masses — uncomfortably low compared with the predicted minimum pancake mass.

The cloud structures could conceivably arise by fragmentation of pancakes. The details of pancake collapse are sensitive to assumptions about the initial distribution of density fluctuations, and only a modest amount of fragmentation is required. But there are more exotic possibilities if gravitinos or other very weakly interacting particles provide the major contribution to the mass density of the universe. One needs particles that are even more weakly interacting than neutrinos: they decouple from the universe at an earlier epoch, when density fluctuations of mass comparable with the horizon size can first grow. In this way, it is possible to have the smallest structures collapse with masses of 10^{13} or even less.

The redshift survey of the Local supercluster, providing information on the velocity field, is also relevant. Aaronson *et al.* (*Astrophys. J.* 258, 64; 1982) combine data on galaxy redshifts with measurements of IR luminosities and neutral hydrogen line profiles. The Tully-Fisher relation, which expresses the empirical correlation between galaxy luminosity and hydrogen profile width, was used to infer the distances to all galaxies in a restricted sample. The authors thus have independently derived distances and velocities, and so are able to develop a model for the velocity field in the Local

supercluster.

They find that the excess in the local density decelerates galaxies relative to the universal expansion. In addition, the Local Group could have a random component of motion relative to nearby galaxies. The result of fitting the data to the model is a value for the deceleration pattern at the position of the Local Group of about 250 km s^{-1} . This sets an important constraint on the mean cosmological density, which cannot exceed a value of about one-third of the critical density for a closed universe. If the mean density were any larger, a correspondingly larger gravitational acceleration would be experienced by the Local Group than is inferred from the observations.

One surprising result has emerged from the analysis of the velocity field in the Local supercluster. Aaronson *et al.* also examined a model which allowed rotation about the supercluster centre in addition to deceleration from the Hubble flow. A significant effect was found: the rotation velocity at the position of the Local Group is $180 (\pm 60) \text{ km s}^{-1}$. It may be premature to overemphasize the significance of such a rotation, but it would be a relatively natural outcome of the pancake theory in which coherent formation of large-scale structure occurs. It is also worth noting that the concept of rotation loses much of its meaning for the Local supercluster: the Local Group has completed only a small fraction of a single orbit since the formation of the supercluster. □

unique in acting even more potently on the κ opiate receptor sub-type, where its actions are only weakly antagonized by naloxone⁷. This and other evidence led Cavkin, James and Goldstein⁷ to suggest that dynorphin might represent a specific endogenous ligand for the κ opioid receptors, but the recent results of Weber *et al.*³ and those described in an accompanying article by Corbett *et al.*⁸ suggest that at least in brain, this role is more likely to be filled by a smaller fragment of the dynorphin molecule, dynorphin₁₋₈.

Using three highly specific radio-immunoassays which recognized dynorphin₁₋₁₇, dynorphin₁₋₈ and α -neoendorphin respectively, Weber *et al.* found that dynorphin₁₋₈ was present in various areas of rat brain in concentrations which were about ten times higher than those of the complete molecule dynorphin₁₋₁₇ and almost precisely equal to those of α -neoendorphin, suggesting that it may share with α -neoendorphin a common biosynthetic origin². Weber *et al.* presented chromatographic evidence to support the claim that the immunoreactive material measured did indeed correspond to dynorphin₁₋₈, which had previously been described as a naturally occurring peptide in extracts of hypothalamus and pituitary^{9,10}. The regional distribution of the dynorphin-related peptides in brain — with the highest concentrations in striatum and hypothalamus — differs from that of the enkephalins, suggesting they occur in distinct neuronal systems.

Some insight into the possible significance of the existence of dynorphin₁₋₈ in brain is provided by the recent results of Corbett *et al.*⁸ on the pharmacological activities of this and various other dynorphin-related peptides. Using a variety of bioassays and radioligand-binding assays capable of discriminating between the μ , δ and κ receptor sub-types, they show a dramatic difference in the pharmacological profiles of Leu-enkephalin and dynorphin₁₋₈ (Leu-enkephalin-Arg-Arg-Ile). Whereas Leu-enkephalin is a selective δ agonist with only weak activity on κ receptors, dynorphin₁₋₈ is less active as a δ agonist but over 1,000 times more potent than Leu-enkephalin as a κ -receptor agonist. The full-sequence peptide, dynorphin₁₋₁₇, is highly potent on all three types of opiate receptor, but less selective than the octapeptide in its κ -receptor activity.

Corbett *et al.* also find that, like most other short opioid peptides, dynorphin₁₋₈ is much more rapidly metabolized in brain

Yet another opioid peptide?

from L. L. Iversen

THE discovery, some seven years ago, that the brain contains its own morphine-like chemicals — the pentapeptides leucine- and methionine-enkephalin¹ — has aroused enormous interest and is rightly regarded as one of the most significant recent advances in our understanding of brain chemistry. Since then, however, a bewildering variety of other naturally occurring opioid peptides has been described². With almost a dozen such compounds already known, the recent report by Weber *et al.*³ of yet another naturally occurring endorphin in brain might at first sight seem merely to add to the difficulties we face in understanding this group of substances. However, the new peptide, which represents the amino-terminal octapeptide fragment of dynorphin (Fig.1), has some unusual properties.

Dynorphin is a potent opioid peptide first identified by Goldstein and colleagues

in pituitary extracts⁴, containing 17 amino acid residues⁵. It belongs to the third branch of the opioid peptide dynasty², of which the first branch comprises the pentapeptide enkephalins and their various slightly elongated forms, and the second branch β -endorphin and other peptides derived from the pro-opiomelanocortin precursor. Dynorphin₁₋₁₇ and related peptides, such as α -neoendorphin⁶ (Fig.1), all contain Leu-enkephalin as their amino-terminal sequence, but dynorphin has strikingly different pharmacological properties from Leu-enkephalin and other opioid peptides. Although, like the other peptides, it is a potent agonist at both μ and δ opiate receptor sub-types, dynorphin is

Dynorphin-related peptides.

Leu-enkephalin	Tyr-Gly-Gly-Phe-Leu
Dynorphin ₁₋₈	Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Ile
Dynorphin ₁₋₁₇	Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Ile-Arg-Pro-Lys-Leu-Lys-Trp-Asp -Asn-Gln
α -Neoendorphin	Tyr-Gly-Gly-Phe-Leu-Arg-Lys-Tyr-Pro-Lys

L. L. Iversen is Director of the MRC Neurochemical Pharmacology Unit, Hills Road, Cambridge CB2 2QH.

than the full-length dynorphin₁₋₁₇ — thus the potency of the octapeptide was markedly increased when a cocktail of peptidase inhibitors was added to the assay system. This is in keeping with a putative transmitter role for dynorphin₁₋₈, whereas the longer peptide, dynorphin₁₋₁₇, may be more important in a hormonal function, perhaps after secretion from the pituitary. The companion peptide, α -neoendorphin, is pharmacologically similar to dynorphin₁₋₈.

Whether these two peptides represent a distinct new family of endogenous ligands for the κ sites in brain and spinal cord remains to be established, but this is certainly an intriguing possibility. The function of these sites remains obscure, but κ -selective drugs, for example ethylketocyclazocine and bremazocine, are known to be effective analgesics and do not substitute for morphine in preventing with-

drawal symptoms in morphine-dependent animals⁹ (they seem to have only low addiction liability). Drugs with such a profile clearly have therapeutic potential, and a better understanding of the κ -receptors and their endogenous ligands is urgently needed. □

1. Hughes, J. *et al.* *Nature* **258**, 577 (1975).
2. Rossier, J. *Nature* **298**, 221 (1982).
3. Weber, E., Evans, C.J. & Barchas, J.D. *Nature* **299**, 77 (1982).
4. Goldstein, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 6666 (1979).
5. Goldstein, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 7219 (1981).
6. Kangawa, K. *et al.* *Biochem. biophys. Res. Commun.* **99**, 871 (1981).
7. Cavkin, C., James, I.F. & Goldstein, A. *Science* **215**, 413 (1982).
8. Corbett, A.D. *et al.* *Nature* **299**, 79 (1982).
9. Minamino, N. *et al.* *Biochem. biophys. Res. Commun.* **95**, 1475 (1980).
10. Seizinger, B., Holtt, V. & Herz, A. *Biochem. biophys. Res. Commun.* **102**, 197 (1981).

Plasma astrophysics at Santa Barbara

from R. Rosner*, E. Zweibel† and V. Trimble‡

At the workshop on 'Space and Astrophysical Plasmas' held this summer at the Institute for Theoretical Physics at the University of California, Santa Barbara, three topics held sway: hydromagnetic shocks and particle acceleration, the interaction of hot and cold plasmas, and hydromagnetic flows. A diverse group of theorists from the laboratory plasma, space plasma and astrophysics communities took part.

The possibility that particles can be accelerated to high energies in the vicinity of collisionless shocks — yielding solar and galactic cosmic rays and relativistic particles in extragalactic radio sources — has received wide attention in recent years. The outstanding problems include the structure of collisionless shocks (the mechanisms which produce a shock transition in collisionless plasmas), the 'injection problem' (how thermal particles can be extracted from the background plasma and accelerated to high energy), the shock efficiency and — most difficult — the interaction between the accelerating particle population and the processes determining shock structure and acceleration efficiency.

These difficult issues can be at least partially addressed by observations of the terrestrial bowshock, produced by the interaction between the solar wind and the terrestrial magnetosphere. E. Greenstadt (TRW) showed that while the shock layer is quite thin when the interplanetary magnetic field is nearly normal to the direction of motion of the shock (quasi-perpendicular case), it is highly irregular and quite extended in space in the quasi-

parallel case (magnetic field nearly parallel to the direction of motion of the shock). A fraction of the incoming solar wind particles is reflected back upstream, and these particles excite low-frequency electromagnetic waves which exert a retarding force on the solar wind. A diffuse energetic ion component is also observed. The terrestrial bowshock (and interplanetary shocks) will evidently provide important 'laboratory' constraints on theoretical studies of shock structure and particle acceleration.

The problems of internal shock structure, including the details of ion dynamics, were addressed by a number of theorists by powerful numerical simulations of collisionless shocks. The simulations show reflected ions and short-wavelength electromagnetic turbulence (D. Forslund, Los Alamos), in reasonable agreement with observations, with ions bent back upstream by the enhanced magnetic field in a postshock magnetic overshoot layer (C. C. Goodrich, University of Maryland). However, the need of an electrostatic potential layer to slow down electrons (in quasiparallel shocks) remains unresolved by the present simulations and observations.

The basic model of shock acceleration, the starting point for discussions at the workshop, is a steady-state model in which

cosmic rays diffuse in the vicinity of a high Mach number quasi-parallel shock which is approximated to by a discontinuity in flow speed. Observations of cosmic-ray secondaries ($E \sim 1$ to 100 GeV) constrain the extent to which cosmic rays can be accelerated to high energies by successive passage through interstellar shocks because the volume in which acceleration occurs does not exceed 10 per cent of the confinement volume (W. I. Axford, Max-Planck Institut). The cosmic rays are scattered by magnetic irregularities, which are assumed to be nearly stationary in the frame of the flow; thus, because of the jump in flow speed across the shock, particles are accelerated by a first-order Fermi process as they scatter back and forth across the shock front, with the well known result that the energetic particle spectrum is quite similar to that of galactic cosmic rays.

Some major issues remain unresolved, however. Thus the efficiency of the shock (the fraction of energy which goes into relativistic particles) is not determined, and the properties of the hypothetical scattering centres, which determine the cosmic ray diffusion coefficient, must be specified *ad hoc*. If the scattering mean-free path is not large compared with the width of the shock transition, the detailed fluid velocity profile must determine the acceleration. But the flow itself is modified by the coupling between the cosmic ray momentum flux and pressure and the thermal gas, the modification being highly sensitive to the details of the coupling (D. Eichler, University of Maryland; H. Volk, Max-Planck Institut/University of Chicago).

There is a very different problem in the transfer of energy between hot and cold plasmas. Thus the structure of the transition layer between the solar chromosphere and corona is crucial to an understanding of the energetics of these plasmas, but the construction of valid model atmospheres is complicated by the high degree of spatial inhomogeneity (including the probable existence of unresolved features), flows in the boundary layer region and the general absence of steady-state conditions (R. Rosner, Harvard University). Even if these difficulties were not present, calculations show that the electron distribution function in the transition region (E. Shoub, Stanford University) is significantly non-Maxwellian. The distribution at a given height is 'contaminated' by superthermal particles representative of the temperature of overlying layers, possibly significantly altering the rates of collisional ionization and excitation (and the heat flux). Thus, the diagnostic problem remains a major difficulty in modelling solar thermal boundary layers.

Different problems arise in heat transport in laser-irradiated plasmas (C. Max, Lawrence Livermore). Experiments show the heat flux often to be much less than predicted by classical transport

*Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138.

†University of Colorado, Boulder, Colorado 80309.

‡University of California, Irvine, California 92717.

formulae. The solution may lie in improved classical calculations; in the suppression of heat flux by plasma instabilities (ion acoustic waves, Weibel instabilities and so on) which increase the electron-scattering frequency; or in the inhibition of thermal conduction by strong (megagauss) d.c. magnetic fields (which occur because the isothermal and isodensity surfaces do not coincide, as in the classic Biermann 'battery' dynamo). In any case, the observed heat flux appears to limit at about 3–9 per cent of the free-streaming value, a result which the simulations do not as yet reproduce.

This observation is clearly relevant to models of heat exchange between tenuous hot plasma and much cooler gas in the interstellar medium of galaxies (discussed by C. McKee, University of California, Berkeley, and S. Balbus, MIT). In such models, the crucial parameters controlling the dynamics are the ratio of the thermal electron mean free path to the dimension of the cool plasma embedded in the hotter gas phase; and the dimensionless coefficient marking the reduction of the heat flux below its free-streaming value. Present calculations have now explored the parameter space of the hot electron-cold plasma interaction in some detail; presumably, the 'zoo' of possible solutions must be constrained by observations (which are likely to lend considerable insight into the dominant mechanisms of energy transport across such steep temperature boundary layers) and by the laboratory laser work (which reduces the arbitrariness of assigning a reduction factor to the saturated heat flux).

The discussions on hydromagnetic (MHD) flows indicated that although much is understood in restricted contexts, and there is an appreciation of the connections on the observational level, an integrated theoretical view of the problem (as is now emerging in the above two cases) is not yet at hand. At the very fundamental level, it is apparent that study of equilibria may be of limited usefulness because, for example, one can now show (E. N. Parker, University of Chicago) that 'most' perturbations lead to departures from equilibrium conditions characterized by some symmetry property (these are indeed the only known equilibria). Furthermore, MHD equilibria which have been shown to be stable to (analytical) perturbations have only remote connections with observed magnetic structures (on, for example, the Sun). It is likely that further progress will require sophisticated numerical (MHD) simulations, as experience in laboratory fusion work has indicated.

There is considerable interest in studying the dynamic interaction between hydrodynamic flows and ambient magnetic fields, such as occurs both within the solar wind and in its interaction with planetary magnetospheres (T. Holzer, HAO); and in the interaction between the relativistic plasma and ambient magnetic fields

surrounding pulsars (J. Arons, University of California, Berkeley). In the case of solar wind streams the origin of the high speeds remains a basic question. Observations have now placed stringent constraints on the available wave momentum flux below the solar corona, which seem severely to constrain the popular Alfvén wave-driven models. Another outstanding problem remains the construction of a self-consistent magnetosphere, a problem which is, as pointed out by Arons, very reminiscent of the global terrestrial magnetosphere problem discussed earlier by J. M. Cornwall (University of California, Los Angeles).

In both cases, local theories encounter the difficulty that the current flow is determined by boundary conditions which, in fact, simply reflect global structure. An interesting further example of hydro-magnetic flows which seem to be ubiquitous are the 'bi-polar' flows, or jets,

discussed by A. Königl (University of California, Berkeley); these are perhaps best known in the context of radio galaxies and relatively exotic objects such as SS433, but are apparently also associated with protostellar objects. The mechanisms responsible for the initial plasma acceleration and for the high degree of collimation are at the heart of current theoretical studies; Königl focused particular attention on radiative driving and on the beaming resulting from the planar symmetry and stratification of gas in accreting protostellar objects.

On the whole, our impression was that this workshop gave substance to the long-rumoured emergence of plasma astrophysics as a well defined discipline; the discussions indicated that a community of scholars has 'gelled' in at least two of the three principal topics of discussion and that the much-discussed, but often in practice neglected, contact with laboratory and space plasma physics has been made. □

Key structures in transposition

from N. Symonds

It is widely agreed that transposition is normally a replicative process but little is known about the location of replication origins within transposons or the events which herald the initiation or termination of DNA synthesis. Of undoubted importance are the sequences at the ends of transposons which remain constant during transposition and, in nearly all cases, are inverted repeats of one another. It was this symmetry in the structure of transposons, together with a considerable amount of genetic data, which led Shapiro¹ to propose his model for transposition which has become the archetype for the numerous alternative schemes subsequently put forward.

The model involves a symmetric intermediate structure possessing two replication forks for DNA synthesis (Fig. 1a). It has been successful in explaining the properties of such transposons as Tn3 and $\gamma \delta$, but it fits less well with the properties of phage Mu from which most of Shapiro's early genetic data were obtained and which, ironically as it turned out later, seems to be the only transposon lacking inverted repeats at its ends.

As well as discrepancies in the genetic data, earlier electron-microscope studies by Harshey and Bukhari² had indicated that the intermediate structures in Mu transposition were not as predicted from the Shapiro model but consisted of 'keys' having circles of variable size attached to

tails of variable length. This led these authors to propose a modification to the Shapiro model with an intermediate structure essentially as depicted in Fig. 1b. The main innovation is that it contains a single replication fork starting from one end only of the transposon: the second end of the target molecule is left unligated until the last stage in the transposition process and in the interim is held by a protein complex which could be the replication complex itself. Protease treatment of such an intermediate then yields a key molecule with a free end.

To obtain definitive evidence that the key structures are indeed transposition

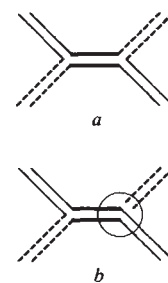


Fig. 1 The initial intermediate structures in transposition according to a, the original model of Shapiro (two replication forks); and b, the modification of Harshey and Bukhari (one replication fork). Solid lines: donor DNA containing the transposon (heavy lines). Dotted lines: target molecule. The circle represents a protein complex and could be located at the replication fork. Normally the donor and target molecules would be on different circular genomes, or at different locations on the same circular genome.

N. Symonds is in the School of Biological Sciences, The University of Sussex, Falmer, Brighton, Sussex BN1 9QG.

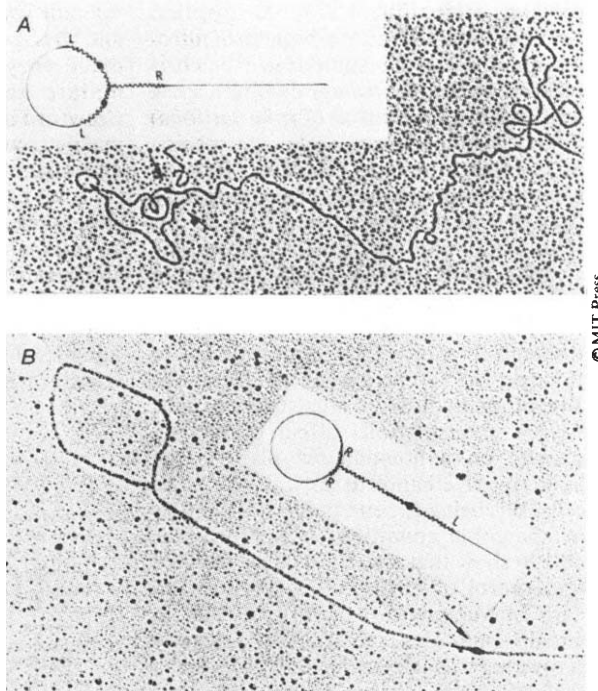
intermediates and to determine from which end of the Mu genome replication is initiated, Harshey and Bukhari³ have recently performed some ingenious experiments using a Mu derivative into which they first engineered a *lac* operator site. After induction, the Mu derivatives were allowed to undergo transposition and the DNA was passed over nitrocellulose filters containing *lac* repressor. All Mu-containing DNA was trapped on the filters by *lac* repressor-operator binding and the *lac* repressor-binding sites were subsequently visualized in the electron microscope by complexing the repressor with ferritin-conjugated antibody (Fig.2). These experiments, and others employing restriction analysis, have shown unequivocally that key structures contain replicating DNA and so are true transposition intermediates. The experiments also throw an unexpected result into the transposition debate: Mu replication can originate with equal facility from either end of the genome.

As the two ends of Mu possess different DNA sequences it is not altogether surprising that there should be some asymmetry in the initial events of Mu transposition. A first guess would be that one end of Mu was involved in DNA initiation, the other in termination; but the new results show unexpectedly that this is clearly not the case. The role of the two ends of Mu (which do possess certain sequence similarities⁴) in transposition must therefore be more subtle. The deletion of either end of Mu abolishes even the introduction of a single-strand nick in the donor molecule, so the two ends of the genome must presumably act in some co-operative way for transposition to be initiated.

If the end of Mu at which replication commences is not determined by the terminal sequences, then other causes must be sought. Two possibilities are the availability of specific enzymes and the DNA sequences at the target sites. The products of two Mu genes, *A* and *B*, are essential for efficient transposition and the *A* product is known to be present in limiting amounts. Perhaps only a single molecule of *A* product is normally available for a particular Mu genome, and it could bind either to the left or the right end, so triggering off replication from either direction.

With regard to the preference of Mu for particular target sites, few sequence data are available. It is usually said that Mu integration occurs completely at random — certainly Mu phage can integrate at a large number of locations in the bacterial chromosome — but it must be remembered that the integration of phage DNA to form lysogens is not necessarily an identical

Fig.2 Electron micrographs of Mu DNA undergoing transposition, showing the key structures as transposition intermediates. The *lac* operator appears as an electron-dense ferritin core; L and R, left and right ends of Mu respectively. In *A*, the ferritin cores are equidistant from the junction between the circle and tail, so replication must have taken place, with the replication fork at the junction. The Mu plasmid is 16.5 kb long and the *lac* operator 5 kb from the left end. As the circle in *A* is 12 kb, the key structure must be a result of replication starting at the left end. In *B*, the circle is 8.9 kb and there is one ferritin core on the tail. The key structure could only have arisen from replication starting at the right end.



process to that which goes on in the experiments considered here, which follow Mu transposition in lysogenic cells which have undergone induction. There is indeed some indirect evidence that in these conditions transposition does occur into non-random sites⁴, and these sites in turn could

determine at which end the replication of Mu commences.

There is one afterthought from this discussion. If Mu can transpose so efficiently without having inverted repeats at its ends, why do other transposons all seem to possess them? □

Survival mechanisms in wetland plants

from Peter D. Moore

AN apparently simple problem faces the wetland ecologist when considering how it is that flood-tolerant plants are able to survive anoxic conditions. The roots of wetland plants are often, sometimes permanently waterlogged and thus subject to low oxygen tensions. Below about eight per cent oxygen saturation (about 320–350 mV at pH 5)¹, ferric ions are replaced by ferrous ions and environmental conditions are distinctly reducing, and yet many species of plant can still survive.

In practical terms, the ecologist encounters two main difficulties. The first is that in anoxic conditions, not only does the plant encounter problems of oxygen supply, but it also experiences a build-up of potentially toxic materials², such as ferrous and manganese ions and perhaps even organic compounds produced by incomplete microbial decay³, which may affect its response. The second is the sheer technical difficulty of producing strictly anoxic conditions in order to test the plant's degree of dependence on oxygen supply.

Field experimentation, such as that of Armstrong and Boatman⁴, has shown that different plants have varying tolerance to anoxia. The wetland grass *Molinia caerulea*, for example, was not found in peats with oxidation-reduction potentials below 50–100 mV (no hydrogen sulphide was present at this oxidation level), whereas the bogbean, *Menyanthes trifoliata*, was able to tolerate lower oxygen tensions and free sulphide. Oxygen undoubtedly diffuses from the above-ground structures into the roots of wetland plants, especially those with anatomical air ducts, such as *Menyanthes*⁵, and may even diffuse out of the root to form an oxygen-enhanced sheath^{6,7}. This can produce complex micropatterns of oxygen concentration in the field, making experimentation difficult.

To overcome the complications due to

Peter D. Moore is a Senior Lecturer in the Department of Plant Sciences, University of London King's College, 68 Half Moon Lane, London SE24 9JF.

1. Shapiro *Proc. natn. Acad. Sci. U.S.A.* 76, 1933 (1979).
2. Harshey & Bukhari *Proc. natn. Acad. Sci. U.S.A.* 78, 1090 (1981).
3. Harshey & Bukhari *Cell* 29, 561 (1982).
4. Kamp & Kahmann *Cold Spring Harb. Symp. quant. Biol.* 45, 329 (1981).

oxygen diffusion, Laing^{8,9} supplied aquatic plants with an atmosphere of nitrogen and showed that some species, such as pickerel weed (*Pontederia cordata*), grew well, whereas reedmace (*Typha latifolia*) did not, again confirming the interspecific differences in anoxia tolerance. But the nitrogen supply used by Laing could not be regarded as totally oxygen free, so his results still did not establish how long various plant species could survive in strictly anoxic conditions.

R.M.M. Crawford has long maintained that the tolerance of a plant species to flooding is a function of its metabolic adaptation to the low oxygen tensions in its tissues. One obvious effect of oxygen scarcity on a non-adapted plant is the build-up of ethanol to toxic levels in its cells, this being the end point of glycolysis in anaerobic conditions. Crawford was able to show that flood tolerance involved the control of ethanol production¹⁰ and with McManmon¹¹ he proposed that in anoxic conditions, the glycolytic pathway is diverted to oxaloacetate and (non-toxic) malate which, if the activity of malic enzyme were impaired, would accumulate and a return to pyruvate and ethanol production would be prevented. In flood-

tolerant species such as *Glyceria maxima* and *Iris pseudacorus*, they found that malic enzyme was indeed inactive and malate accumulated. Subsequently, Crawford has shown that in tolerant trees, malate accumulates in response to flooding and ethanol production is controlled¹², and that in pea seedlings anoxic death is closely related to the build-up of ethanol, an internal concentration of 60 μ M being apparently the critical survival threshold¹³.

If the response of a plant species to flooding indeed depends on its metabolic adaptations rather than its efficiency in translocating oxygen from another part of

its body, then one should be able to observe interspecific differences in survival under strictly anoxic conditions. It is precisely this experiment that Barclay and Crawford¹⁴ have now carried out to vindicate the metabolic theory. They have grown a range of species under an oxygen-free nitrogen atmosphere (85 per cent N₂, 10 per cent H₂ and 5 per cent CO₂) in the presence of a palladium catalyst to ensure a total absence of oxygen. They found that some species, for example the bulrushes *Scirpus lacustris* and *S. maritimus*, not only survived but continued to grow over 14 days, whereas others, such as *Phragmites* and *Iris*, survived but showed no growth after 7 days, and another group, for example *Glyceria maxima* and *Juncus effusus*, failed to survive the week. This latter group included, rather surprisingly, rice (*Oryza sativa*) which has previously been regarded as particularly anoxia tolerant.

The capacity of some wetland plants to survive in conditions of strict anoxia, together with the degree of variation found among the species tested, strengthen very considerably the assertion that flood tolerance can, at least in part, be explained at the cellular, metabolic level. □

1. Pearsall, W.H. *J. Ecol.* 26, 298 (1938).
2. Rutter, A.J. *J. Ecol.* 43, 507 (1955).
3. Drew, M.C. *Curr. Adv. Pl. Sci.* 11, 1 (1979).
4. Armstrong, W. & Boatman, D.J. *J. Ecol.* 55, 101 (1967).
5. Coult, D.A. & Vallance, K.B. *J. exp. Bot.* 9, 384 (1958).
6. Jeffery, J.W.O. *J. Soil Sci.* 12, 172 (1961).
7. Jeffery, J.W.O. *J. Soil Sci.* 12, 317 (1961).
8. Laing, H.E. *Am. J. Bot.* 27, 574 (1940).
9. Laing, H.E. *Bot. Gazette* 102, 712 (1941).
10. Crawford, R.M.M. *J. Ecol.* 54, 403 (1966).
11. McManmon, M. & Crawford, R.M.M. *New Phytol.* 70, 299 (1971).
12. Crawford, R.M.M. & Baines, M.A. *New Phytol.* 79, 519 (1977).
13. Barclay, A.M. & Crawford, R.M.M. *J. exp. Bot.* 32, 943 (1981).
14. Barclay, A.M. & Crawford, R.M.M. *J. exp. Bot.* 33, 541 (1982).

100 years ago

NEW OR RARE ANIMALS IN THE ZOOLOGICAL SOCIETY'S LIVING COLLECTION

THE PIGMY HOG (*Porcula salvania*). — Few additions to the Zoological Society's living collection of late years have attracted more attention than the Pigmy Hogs of Nepal, of which the first specimens ever imported into Europe reached the Gardens in May last.

For our first knowledge of the existence of this diminutive form of the pig-family in the sub-Himalayan forests we are indebted to the researches of Mr. Bryan H. Hodgson, formerly Resident at the Court of Nepal, who described the Pigmy Hog so long ago as 1847, in an article published in the *Journal of the Asiatic Society of Bengal*. While the Wild Boar, or a species closely resembling it abounds all over India, the Pigmy Hog is exclusively confined, as Mr. Hodgson tells us, to the deep recesses of the primeval forests of the Terai of Nepal and Bhotan, where it roams about in herds. It is very rarely seen even by the natives. A well-known hunter informed Mr. Hodgson that during fifty years' abode in the Saul forests he had obtained but three or four of these animals to eat, partly owing to their scarcity, and partly to the speed with which the females and young disperse, and to the extraordinary vigour and activity with which the males defend themselves while their families are retreating.

THE CABOT'S TRAGOPAN (*Cerionis caboti*). — The Tragopans, or Horned Pheasants, constituting the genus *Cerionis* of naturalists, must be ranked amongst the finest and most brilliantly coloured representatives of the splendid group of Indian game birds. Two of them — the Crimson Tragopan of the Central and Eastern Himalayas, and the Black-headed Tragopan of the Western Himalayas and Cashmere, are well known to Indian Sportsmen, and are familiar objects of pursuit, though we believe, by no means easily procured. The Crimson Tragopan was introduced into Europe by the Zoological Society in 1859, and has frequently bred in their Gardens, as has likewise the Temminck's Tragopan (*Cerionis temmincki*), first received by the Society in 1864.

Between the furthest known eastern range of the Crimson Tragopan and the frontiers of China a fourth species of *Cerionis* has its home — Blyth's Tragopan (*C. blythi*).

The fifth and last species of Tragopan, lately acquired by the Zoological Society, is still more rare and little known than the four above-mentioned members of the genus. Cabot's Tragopan, as it is called, was described in 1857 by the late Mr. Gould, and subsequently figured in his great illustrated work on the Birds of Asia. Its habitat is South-Eastern China, but little is yet known of its exact range. The only naturalist who has met with it in its native wilds is the celebrated Chinese explorer, M. le Père David. M. David, in his "Oiseaux de la Chine," tells us that he found this fine Gallinaceous bird rather common in the wooded mountainous range which separates the provinces of Folien and Kiangsi, when he traversed this district in the autumn of 1873.

From *Nature* 26, 603, 19 October 1882.



FIG. 25.—The Cabot's Tragopan.

REVIEW ARTICLE

Postnatal development of the visual cortex and the influence of environment

Torsten N. Wiesel*

The following is the lecture delivered by the author in Stockholm on 8 December 1981 when he received the Nobel Prize in Medicine, which he shared with Roger Sperry and David H. Hubel. The article is published here with permission from the Nobel Foundation and will also be included in the complete volume of Les Prix Nobel en 1981 as well as in the series Nobel Lectures (in English) published by Elsevier.

In the early 'sixties, having begun to describe the physiology of cells in the visual cortex of the adult cat¹, David Hubel and I decided to investigate how the highly specific response properties of cortical cells emerged during postnatal development. We were also interested in examining the role of visual experience in normal development, a question raised and discussed by philosophers since the time of Descartes. The design of these experiments was undoubtedly influenced by the observation that children with congenital cataracts still have substantial and often permanent visual deficits after removal of the cataract and proper refraction². Also, behavioural studies had shown that animals raised in the dark or in an environment devoid of contours have a similar impairment of their visual functions^{3,4}.

Because of the difficulties associated with raising kittens in total darkness, we decided to fuse the lids by suture. This procedure prevented any form vision without completely depriving the animal of light. We expected this to be an effective procedure because cortical cells respond to contours and are insensitive to changes in levels of diffuse light⁵. Initially we raised kittens with only one eye closed, using the other eye as a control. This design turned out to be fortunate because the effects of single eye closure on the visual cortex are more dramatic than the results obtained from animals raised with both eyes occluded or kept in the dark.

Our initial findings were that kittens with an eye occluded by lid suture during the first 3 months of life were blind in the deprived eye and that in the striate cortex the majority of the cells responded only to stimulation of the normal eye⁶. This defect seemed to be localized to the visual cortex, perhaps at the site of interaction between geniculate afferents and cortical cells⁷. From another series of experiments, we found that the properties of orientation specificity and binocularity developed through innate mechanisms⁸. This result, taken together with the monocular deprivation experiment, indicated that neural connections present early in life can be modified by visual experience. Such neural plasticity was not observed in the adult cat, but existed only during the first 3 postnatal months⁹.

The early experiments were done in the cat, but we soon turned our attention to the rhesus monkey. Having demonstrated that cells in the monkey visual cortex also respond selectively to lines of different orientations and often are binocular¹⁰, we showed that the monkey was also susceptible to visual deprivation¹¹, a finding subsequently confirmed and extended¹²⁻¹⁵. Further advances in our understanding of the nature of and mechanism underlying the deprivation phenomena depended on working out some of the functional architecture of the visual cortex. This was done through further physiological experiments in the normal animal and by using

newly developed anatomical methods¹⁶⁻²¹. Over the years, our parallel pursuit of the normal and developmental studies has accelerated our progress in both areas. For example, while the deprivation experiments depended on the understanding of the functional architecture of the normal adult animal, we were alerted to the existence of ocular dominance columns in the cat by experiments we had done in strabismic animals²².

In this lecture I will present our current understanding of the development of the monkey visual cortex and the role of visual experience in influencing neural connections. Rather than attempt to discuss in any detail the now extensive literature in the field, I will emphasize the work carried out in our laboratory²³⁻²⁵. Hubel and I did much of this work in collaboration with Simon LeVay.

Monocular deprivation

The procedure of suturing a monkey's eyelid shut creates a condition similar to a cataract, since though the light reaching the retina through the closed lid is only slightly attenuated (by a factor of 3), the forms of objects are no longer visible. As mentioned above, when the deprived eye is opened after months of deprivation, the animal is unable to see with it: there are no obvious changes in ocular media, the retina or the lateral geniculate nucleus (LGN) that can explain this deficit; instead marked changes have occurred at the level of the primary visual cortex (striate cortex). Even if the ocular media are clear, the occluded eye develops with time a marked axial-length myopia (up to 12 dioptres over a 1-yr period)²⁶.

One way of seeing the change is to record from cells in the striate cortex and determine their ocular preference⁷. In the monkey, there is normally a fairly even balance between cells driven preferentially by one eye and cells driven preferentially by the other¹⁰. In layer 4, most cells are strictly monocular, and outside layer 4 they are usually binocular, although they still tend to respond more strongly to stimulation of one eye than the other. About as many cells prefer stimulation of the left eye as prefer stimulation of the right (Fig. 1a). In conditions of monocular deprivation, however, the great majority of cortical cells are driven exclusively by the nondeprived eye (Fig. 1b)¹¹⁻¹³. One could ask whether this can be accounted for in terms of changes occurring at the level of the LGN. Although the cells lying in the geniculate layers that receive input from the deprived eye are smaller than those in the nondeprived layers (Fig. 2), they are present in normal numbers, respond briskly to stimulation of the deprived eye, and have normal receptive fields. Since geniculate cells are functionally normal and the cortical cells are altered in their properties, there must be some change in the effectiveness of the geniculocortical connections. We were interested in investigating whether there were any structural changes associated with this abnormality.

The first aspect of cortical organization to be examined is the pattern of input of the geniculate afferents to the cortex.

* Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA.

This can be done with the autoradiographic technique for tracing neuronal connections, either transynaptically from the eye^{18,21} or by a fibre stain¹⁹. When in the normal monkey the input from the LGN reaches layer 4 of the cortex, the information from the two eyes is still segregated. The input from each eye is distributed into a series of branching and anastomosing bands about 0.5 mm wide, which alternate with similar bands serving the other eye (Fig. 3a). This pattern of innervation forms the anatomical basis for ocular dominance columns. Cells in the superficial and deep layers, while tending to be more binocular than cells in layer 4, are still more strongly influenced by the eye that provides input to the column in which they reside. The relative influence of the two eyes is shown by making tangential electrode penetrations through different cortical layers. Such penetrations in the normal monkey show regular changes in eye preference as expected from the columnar arrangements (Figs 3a, 4a).

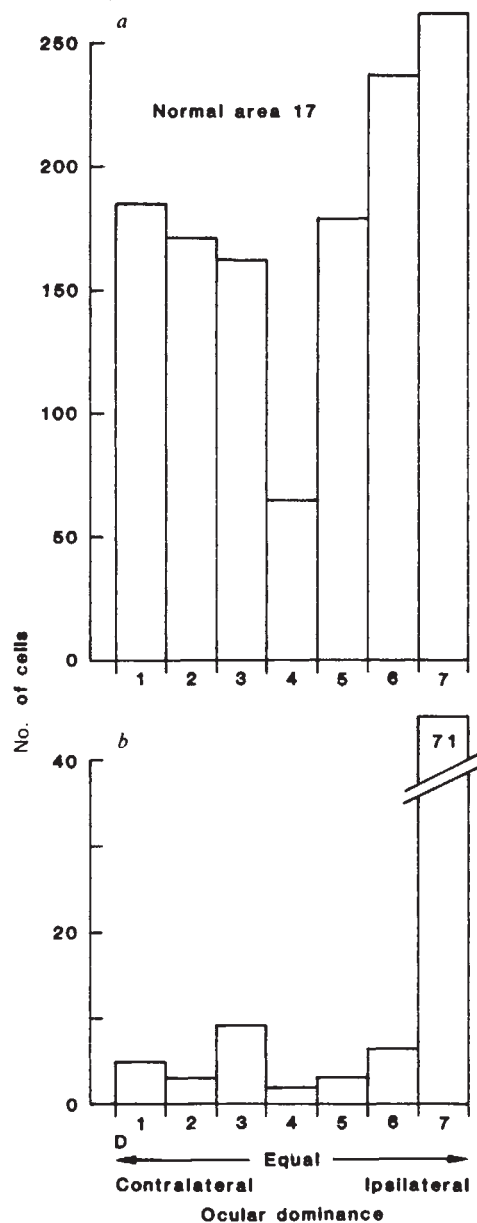


Fig. 1 Ocular dominance histograms in normal and monocularly deprived rhesus macaque monkeys. *a*, Cells (1,256) recorded from area 17 in normal adult or juvenile rhesus monkeys¹³. *b*, Histogram obtained from a monkey in which the right eye was closed at the age of 2 weeks for 18 months¹⁵, showing the relative eye preference of 100 cells recorded from the left hemisphere. The letter D indicates the side of the histogram corresponding to dominance by the deprived eye. Cells in layer 4C are excluded here and in other histograms. Cells in group 1 are driven exclusively from the contralateral eye, those in group 7 from the ipsilateral eye, those in group 4 are equally influenced, and the remaining groups are intermediate.

In an animal that has undergone monocular deprivation, the geniculate terminals with input from the nondeprived eye take over much of the space that would normally have been occupied by terminals from the deprived eye (Fig. 3b)^{13,15}. The input from the deprived eye has shrunk down to occupy the small strips lying between the terminals of the input to the nondeprived eye. Tangential electrode penetrations through cortical layers reveal long expanses of cells driven by the nondeprived eye interrupted by small patches of cells that are either unresponsive or driven by the deprived eye (Fig. 4b). As will be shown below, this expansion of the input from the nondeprived eye occurs at the level of single geniculate afferents. Cells in the deprived layers of the LGN are smaller than normal. One reason for this is that their shrunken cortical arbors may require a smaller soma to maintain them, as originally proposed by Guillery and Stelzner²⁷.

Morphological examination of the LGN in these animals showed a good relationship ($r = 0.91$) between the relative size of normal and deprived cells and the relative size of normal and deprived ocular dominance columns in layer 4C¹⁵. Thus, measuring geniculate cell sizes is yet another means of evaluating the effects of monocular closure.

From the histogram shown in Fig. 1 one cannot tell whether many cells have changed allegiance from the deprived to the nondeprived eye or have simply become unresponsive. The autoradiographic labelling of the afferents in layer 4 (Fig. 3b) shows that a greater proportion of the cells receive direct input from the nondeprived eye. The consequence of this change is that cells at later stages have shifted allegiance from the deprived to the nondeprived eye, rather than become unresponsive. This conclusion is supported by the physiological findings that the large majority of cells in superficial and deep layers respond only to the stimulation of the normal eye (Fig. 4b).

The critical period

Having observed these dramatic effects of monocular suture early in an animal's life, we wanted to determine if there was a period over which the cortex retained its plasticity.

Our experiments in adult cats and monkeys^{6,15} showed that long periods of monocular lid suture did not result in the sort

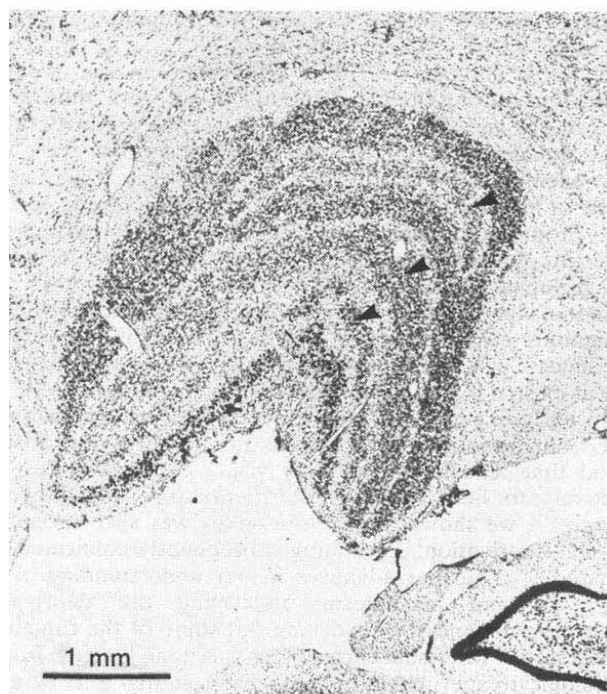


Fig. 2 Frozen coronal section through the right lateral geniculate nucleus of the monkey with right eye closed at 2 weeks for 18 months (Fig. 1b). Atrophy in the layers receiving input from the deprived eye is indicated by arrows. Stained with cresyl violet (from ref. 13).

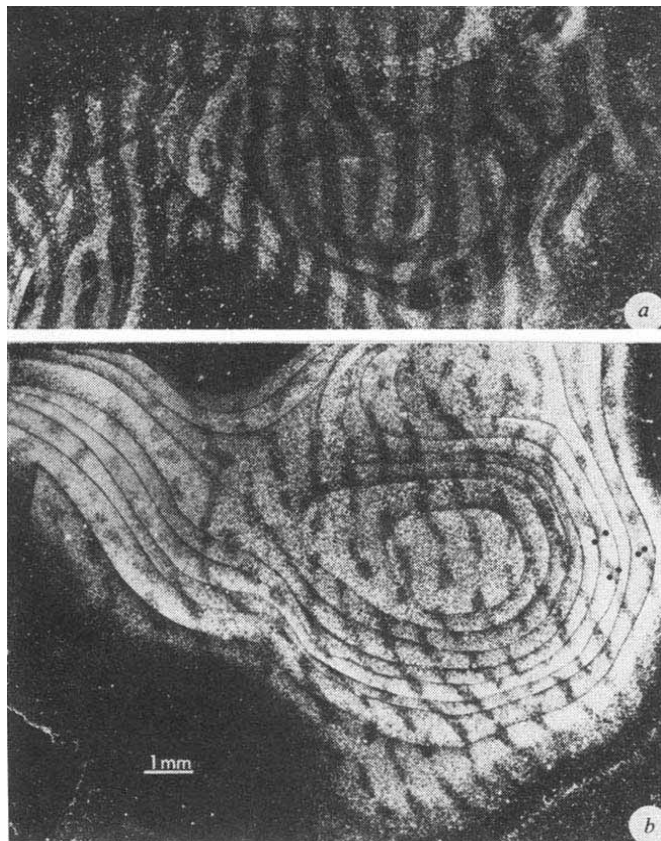


Fig. 3 Dark-field autoradiographs of monkey striate cortex 2 weeks after the injection of ^3H -proline in the vitreous of one eye. *a*, Normal monkey, a montage of a series of tangential sections through layer 4C. The light stripes, representing the labelled eye columns, are separated by gaps of the same width representing the other eye. *b*, Monocularly deprived monkey, again a montage from a series of tangential sections through layer 4C. This is the same monkey shown in Figs 1*a* and 2, which had the right eye closed at 2 weeks for 18 months. The input from the normal eye is in the form of expanded bands which in places coalesce, obliterating the narrow gaps which represent the columns connected to the closed eye.

of changes in the visual cortex described above. Instead, we found that there is a definite period of time, early in life, during which the visual system shows this lability. We termed this the 'critical period'. The permanent visual deficits observed in children with congenital cataracts are therefore most probably a result of changes in the visual cortex that occurred during the critical period. Adult humans suffering from cataracts for many years will have normal vision when the cataracts have been removed, presumably because they are well past their critical period at the onset of the disease.

The critical period in the monkey was estimated by closing one eye at different ages and keeping it closed for several months or longer¹⁵. The deprivation effect was gauged by the relative influence of the two eyes on single cortical cells (ocular dominance distribution), by the distribution of the input from the two eyes in layer 4 (with the autoradiographic technique shown in Fig. 3), and by a comparison of the cell sizes in deprived and nondeprived layers of the LGN. The physiological results in monkeys with one eye closed at 2 weeks of age, 10 weeks, 1 yr and in the adult are illustrated in terms of ocular dominance histograms in Fig. 5. The earliest closure produced the most severe shift of preference towards the normal eye. The same degree of shift could be seen up to an age of 6 weeks. At that age the animal's susceptibility to monocular deprivation began to decline, but it was still pronounced at 10 weeks and was detectable at 1 yr. There were no cortical changes when the closure was done in the adult.

The changes occurring in the geniculocortical innervation were in general agreement with the physiology, though the time course was different (Fig. 6). Animals with a closure at 2 and

5½ weeks showed the expected expansion of the nondeprived geniculate terminals; closure at 10 weeks showed a more moderate expansion, and at 1 yr the pattern was indistinguishable from that in the adult. Geniculate cell sizes in the deprived layers changed in a parallel fashion, showing marked shrinkage at early closures, moderate reduction in closure at 10 weeks, and no change when closed at 1 yr. Since in the closure at 1 yr we observed physiological changes in the absence of a change in the pattern of geniculate innervation, there must presumably be changes at subsequent levels in the cortical circuit^{13,14}. At any rate, in the adult even this higher level of plasticity disappears.

The high degree of susceptibility to deprivation at early ages is also apparent from experiments in which one eye in monkeys was closed for short periods. Before 6 weeks of age, it was sufficient to close an eye for a few days to obtain substantial change in eye preference. The ocular dominance histogram from a monkey with one eye closed for 12 days is shown in Fig. 7*a*. During the subsequent several months a marked change required several weeks of closure, and during the second year any change required months of closure.

From these and similar experiments by us^{13,15} and others^{14,28}, we conclude that the macaque monkey is highly susceptible to monocular deprivation during the first 6 weeks of life, at which age the sensitivity declines progressively, so that at 1½ to 2 years the monkey loses this type of neural plasticity. The length of the critical period varies among species. In cats it is 3–4 months^{9,29}, and clinical observations in humans suggest that it may extend to 5–10 yr, though the susceptibility to deprivation seems most pronounced during the first year and declines with age^{30–32}.

Recovery from deprivation

Monocular closure during the entire critical period in cats and monkeys leads to permanent blindness (refs 33, 34 and T.N.W.,

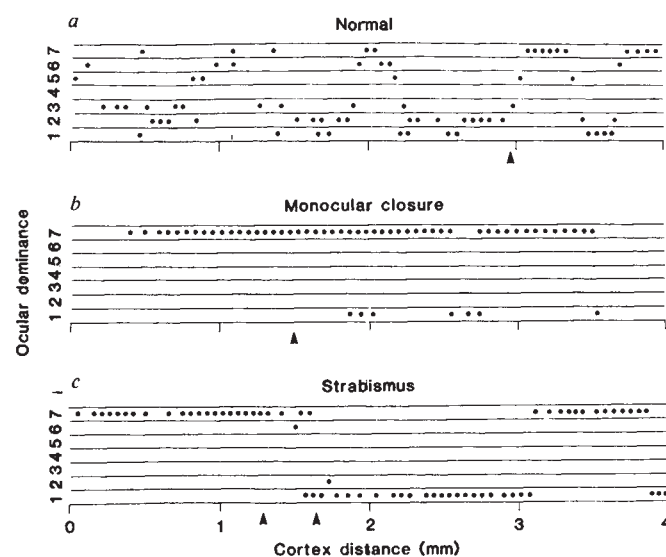


Fig. 4 Eye preference of cells recorded in oblique penetrations through the cortex. Ocular dominance categories 1 to 7 are shown relative to the distance the electrode penetrated through the cortex. *a*, Penetration in a normal monkey showing the sinusoidal shift in eye preference with distance. The arrowhead indicates that the electrode entered layer 4C, in which cells are monocular, and there are abrupt shifts of dominance from one eye to the other. *b*, Oblique penetration in a monkey raised with monocular closure (same monkey as in Figs 1*b*, 2, 3*b*). Outside layer 4, all cells are driven only by the normal eye, and in layer 4C (arrow) there are only short stretches of cells with input from the deprived eye. The overlap of input from left and right eye is not present in the normal monkey. *c*, Oblique penetration in a monkey with a 10° convergent squint produced by sectioning the lateral rectus at 3 weeks (same animal as in Fig. 11*a*). The penetration was made through the striate cortex when the animal was 3½ yr old. Even outside of layer 4, cells were monocular, with equal stretches of cortex dominated by either eye.

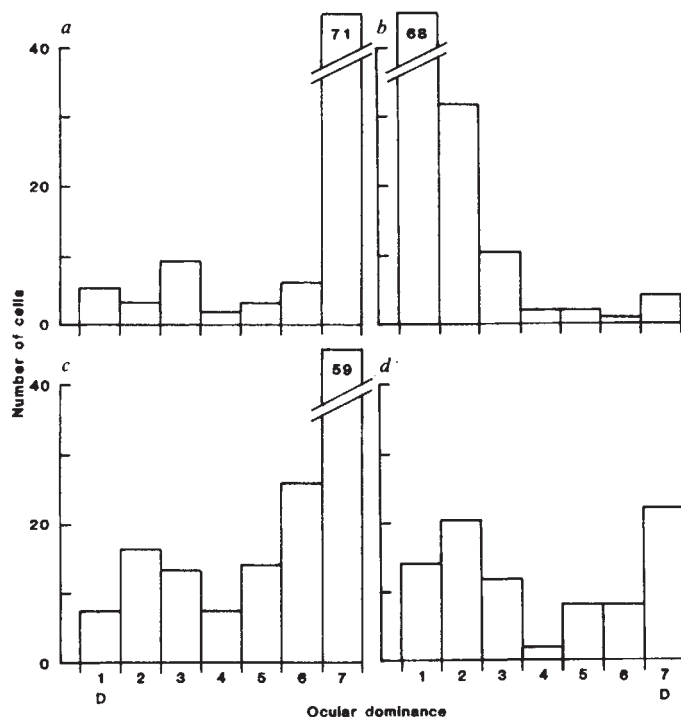


Fig. 5 Ocular dominance histograms of monkeys with one eye occluded by lid suture at different ages and examined after relatively long periods of closure¹⁵. *a*, Same monkey as in Figs 1*b*, 2, 3*b*, 4*b*. The right eye was closed at 2 weeks for 18 months. *b*, A monkey with right eye closed at 10 weeks for 4 months. The strong dominance of normal eye was not as pronounced as at earlier closure. The duration of deprivation was relatively short, but our experience is that at this age the main changes in eye preference occur within the first few months of closure. *c*, A monkey with right eye closed at 1 yr for a period of 1 yr. Preference shifted moderately towards the nondeprived eye, particularly for cells in layers 2 and 3. *d*, Adult monkey (6 yr old) with one eye occluded for 1½ yr. There was no obvious difference in eye preference from that observed in the normal monkey.

unpublished). Presumably there is no recovery of vision after the eye is opened because the pattern of geniculate innervation and the eye preference of cortical cells can no longer be modified. During the period of high susceptibility, partial recovery of vision in the deprived eye is possible after brief periods of monocular closure (refs 15, 35, 36 and T.N.W., unpublished). This was shown in a monkey with one eye closed between days 21 and 30, after which the monkey lived with both eyes open for a period of 4 yr. Initially the animal appeared blind in the deprived eye, but with time it slowly regained the use of the eye; the final acuity in that eye was 20/80 to 100 and in the nondeprived eye, 20/40. The recordings from the striate cortex showed a marked dominance of the nondeprived eye (Fig. 4, right). If there was an increase in the number of cells driven by the once-deprived eye, it was not very obvious. There was a marked narrowing of the deprived columns and a corresponding widening of the nondeprived ones. Thus even if there had been some behavioural recovery, these results demonstrate that a few days of monocular closure had caused clear physiological and anatomical changes in the striate cortex.

These results are relevant to observations in children who have been monocularly deprived for short periods of time. When tested later, some children were found to have reduced acuity in the once-patched eye, the degree of deficit dependent on how young the child was at the time of patching^{37,38}. The experience in children with cataract removal indicates that surgery must be performed very early in the critical period in order to prevent the appearance of any deficit^{39,40}.

A procedure commonly used in children with strabismic amblyopia is to place a patch over the good eye to improve vision in a weak eye. In monocularly deprived animals it was possible to open the sutured eye and close the normal eye⁹, here termed 'reverse suture'. Both in the cat and monkey, reverse suture led to a complete switch in eye preference if it was done within the early part of the critical period^{15,28,29,41}. The geniculate innervation of layer 4C also reversed so that the shrunken regions controlled by the initially closed eye expanded at the expense of the other eye, and consequently the cortical cells switched eye preference in favour of the eye closed first^{15,41,42}. An example is shown in Fig. 8*a*, in which the

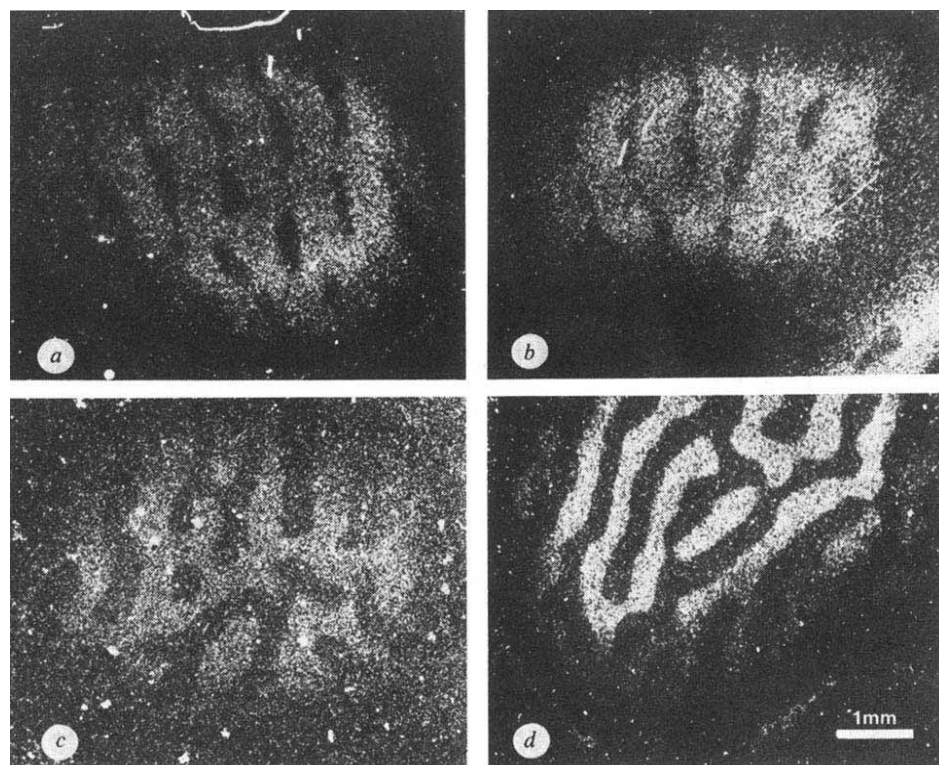


Fig. 6 Autoradiographic labelling patterns from the striate cortex of four monocularly deprived monkeys illustrating the distribution of geniculate terminals in layer 4C after closures at different ages. In all cases the normal (left) eye was injected with ³H-proline, thereby labelling the nondeprived geniculate terminals (from ref. 15). *a*, Right eye closed at 2 weeks for 18 months. Same animal as in Figs 1*b*, 4*b*, 5*a*. *b*, Right eye closed at 5½ weeks for 16 months. *c*, Right eye closed at 10 weeks for 4 months. Same animal as in Fig. 5*b*. *d*, Right eye closed at 14 months for 14 months.

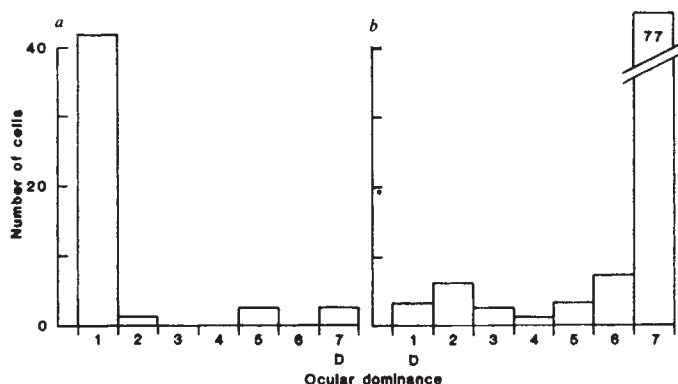


Fig. 7 Ocular dominance histogram for 47 neurones recorded in a 20-day-old monkey whose right eye had been closed since 8 days of age. The physiological picture is similar to that seen after months of deprivation (from ref. 15). *b*, Ocular dominance histogram of 99 cells recorded in a monkey whose right eye was closed from 21 to 30 days of age. Despite a subsequent 4 yr of binocular vision, most cortical neurones were still unresponsive to stimulation of the right eye (from ref. 15).

eye reversal was done at 3 weeks and the recordings 8 months later. The ocular dominance histogram shows that the initially closed eye, which at the time of eye reversal would have influenced very few cortical cells, became strongly dominant. The autoradiography (Fig. 9a) shows a marked expansion in layer 4C β of the initially deprived geniculate terminals. When reversal was done at 6 weeks, the physiology indicated a complete reversal, with a strong dominance of the initially deprived eye (Fig. 8b). Such a marked shift was not reflected in the innervation of layer 4, in which the initially deprived eye had succeeded only in regaining its normal territory (Fig. 9c). This indicates that a significant part of the changes were occurring at the level of intrinsic cortical connections. Finally, reversing at 1 yr failed to restore any of the function of the initially deprived eye (Figs 8c, 9d). Although it is possible to cause changes by monocular deprivation at 1 yr (Fig. 5c), it appears to be more difficult to repair connections that have already been changed once.

Looking more closely at the autoradiography of the geniculate input to layer 4C in the monkey with the reversal at 3 weeks (Fig. 9a, b), one sees a surprising result. The initially deprived eye took over much of the area of innervation of the lower part of layer 4C (4C β), but failed to reverse the dominance of the other eye in the upper part of layer 4C (4C α). Apparently the eye preference of the majority of cortical cells is determined primarily by the cells in layer 4C β (Figs 8a, and 9a, b). Layer 4C β is innervated by cells in the dorsal part of the LGN (parvocellular layers), and layer 4C α by cells in the ventral part of the same nucleus (magnocellular layers). The result of the eye-reversal experiment indicates that the critical period is different for the two cell types. Whereas the critical period is over for the magnocellular input at 3 weeks, the parvocellular input apparently begins to lose its ability to expand at 6 weeks (Fig. 9c), a time when intracortical connections still show considerable plasticity. This result suggests that throughout the brain each functional unit has a unique programme of development.

Mechanism: disuse versus competition

These experiments demonstrate that when a binocular cortical cell is not stimulated by a given eye, the input from that eye

drops out. Other forms of visual deprivation have shed some light on the mechanism of the effect of monocular deprivation. For example, if disuse were an important factor, one might expect that with both eyes closed, cortical cells would not be driven by either eye. Experiments in cats and monkeys raised in conditions of binocular deprivation showed, however, that cells were readily driven by the two eyes^{29,43,44}. The cortex in the monkey was nonetheless altered substantially, in that very few cells were binocularly responsive⁴⁴. This is illustrated in Fig. 10, in which a monkey had both eyes sutured from birth to 4 weeks of age. Except for the obvious lack of binocular cells, the cortex seemed quite normal. The cells were briskly responsive, showed a high degree of orientation selectivity, and had regular sequences of shifts in orientation preference. From tangential electrode penetrations we were also able to see a

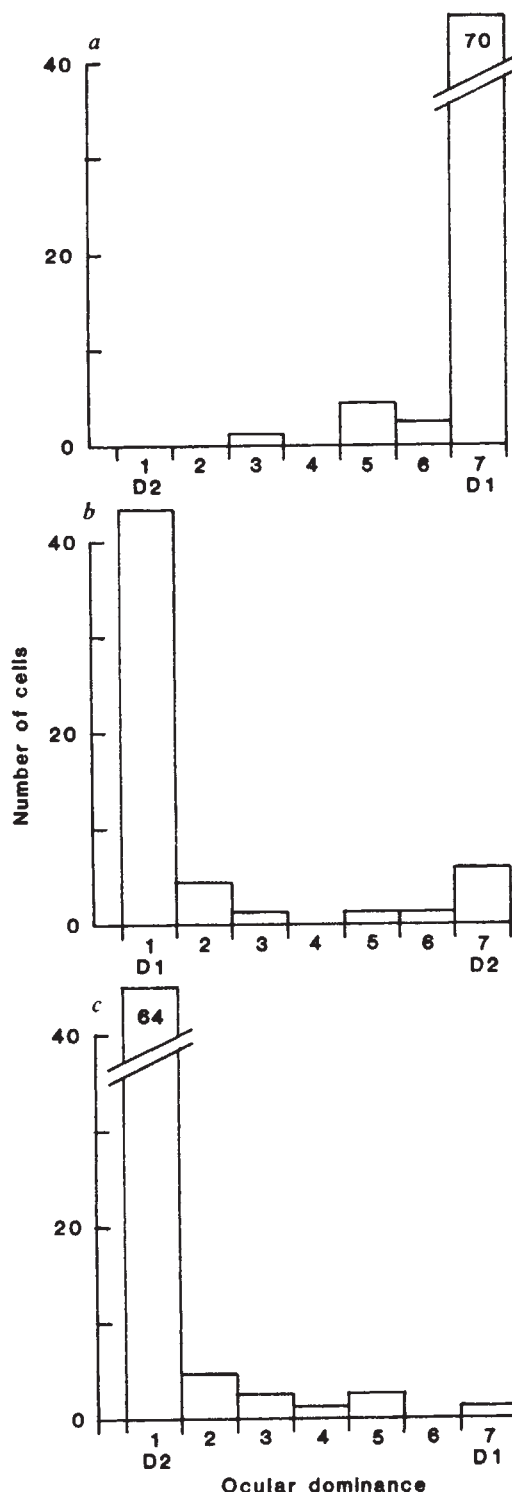
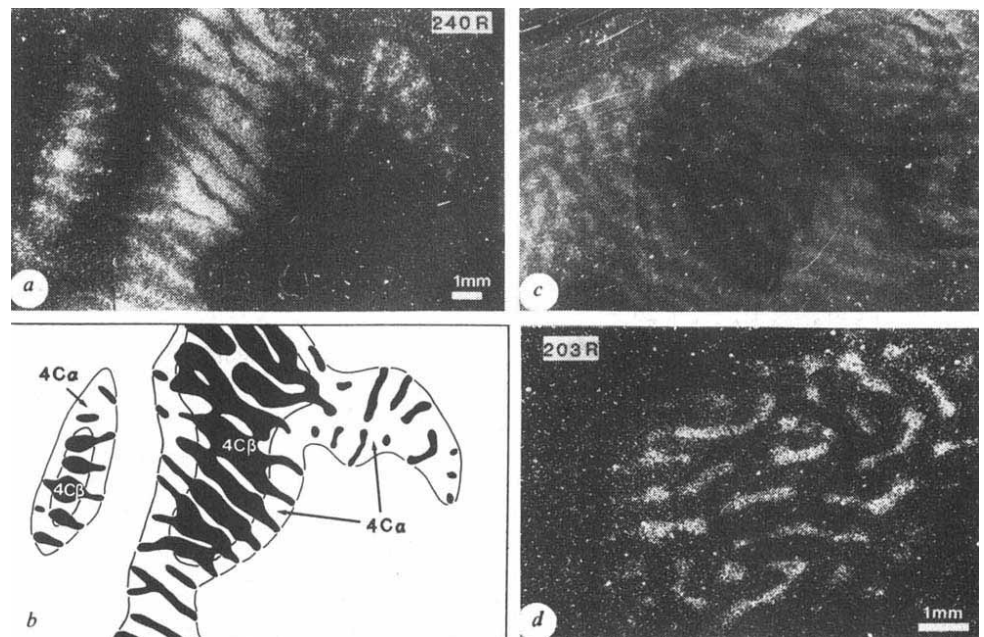


Fig. 8 (Right) Ocular dominance of cortical cells recorded in three monkeys in which reversed suture was done at various ages. *a*, Cells ($n = 77$) recorded from the right striate cortex of a monkey with the right eye (D_1) closed at 2 days for 3 weeks and the left eye (D_2) closed at 3 weeks for about 8 months. Nearly all neurones responded only to the initially deprived right eye (D_1). *b*, Neurones ($n = 56$) recorded from the left striate cortex. The right eye (D_1) was closed at 3 days for 6 weeks and the left eye (D_2) at 6 weeks for 4½ months. Again, nearly all neurones were driven exclusively by the initially deprived right eye (D_1). *c*, Neurones ($n = 74$) recorded from the left striate cortex. The right eye (D_1) was closed at 7 days for 1 yr, and the left eye (D_2) at 1 yr for 2½ yr. In this case the reversal had no effect: nearly all the cells responded only to the initially open eye (from ref. 15).

Fig. 9 Effect of reversed suture at various ages on the labelling pattern of geniculate terminals in layer 4C. *a*, Same monkey as in Fig. 8*a*, in which reversed suture was done at 3 weeks of age. The initially deprived (right) eye was injected with ^3H -proline. In the central region of this single tangential section, the labelled bands are expanded. In the surrounding belt the bands are contracted. These two regions correspond to the β and α sublaminae of layer 4C. *b*, Key to *a*, showing the distribution of label (in black) and the boundaries of the sublaminae 4C α and 4C β , traced from an adjacent section stained with cresyl violet. The thin labelled bands in 4C α run into the centres of the enlarged bands in 4C β , meaning that the two sets of bands, although differing in width, are still in register with each other, as they are in normal animals. *c*, The autoradiographic montage of the labelling pattern in a monkey with reversed suture at 6 weeks (compare with Fig. 8*b*). The initially deprived right eye was injected. The labelled bands are of about normal width, indicating a recovery from the effects of the early deprivation, but not a complete reversal. Most of the montage shows layer 4C β . In layer 4C α , the labelled columns remained shrunken (not shown). *d*, Autoradiographic montage from monkey with reverse suture at 1 yr of age (Fig. 8*c*). The labelled columns (for the initially deprived right eye) remain shrunken, indicating that the late reversal did not permit any anatomical recovery.



clear segregation of the cells into unusually distinct ocular dominance columns, even outside layer 4. When monkeys are kept in the binocularly deprived condition for many months, a considerable fraction of the cells are unresponsive or respond only sluggishly, and they often lack orientation preference. Binocular closures in kittens had a similar effect except that neither short- nor long-term deprivation led to an obvious loss of binocular cells^{29,43}.

Evidence for competitive mechanisms has also been found by measurements of geniculate cell sizes in an ingenious set of experiments^{27,45,46}. First, it was demonstrated in kittens with monocular occlusions that deprived cells in the monocular segment of the nucleus were of normal size, whereas those in the binocular segment showed marked shrinkage²⁷. Next, Guillery produced a monocular region in the zone of binocular overlap by making a local retinal lesion in the normal eye of monocularly occluded kittens⁴⁵. Again the deprived geniculate cells with no competitive input from the other eye were of normal size, and those outside the topographical area corresponding to the retinal lesion showed the usual shrinkage. Finally, binocular closure in kittens apparently did not reduce geniculate cell size⁴⁶, as we originally reported⁴³. These experiments support the hypothesis that competitive mechanisms rather than disuse are prime factors in producing the changes observed in conditions of monocular deprivation.

Because many cortical cells are binocular from birth, the loss in the monkey of binocular cells at early times after closure suggests that in order for cortical cells to sustain a binocular input the two eyes must work together. Another situation that interrupts coordinated activity from the two eyes is strabismus⁴⁷. One way of producing experimental strabismus is to section an extraocular muscle. Sectioning the lateral rectus causes the eye to deviate inwards (convergent strabismus), whereas sectioning the medial rectus produces an outward deviation of the eye (divergent strabismus). After surgery, the sectioned muscle usually reattaches behind the original site, so that, except for the misalignment, normal eye movements are restored. In four monkeys with convergent strabismus, the operation was performed between 3 and 5 weeks (T.N.W. and D. H. Hubel, unpublished). When the animals were examined after a year or more, three of them had normal acuity in both eyes but lacked the ability to fuse the images in the two eyes. The striate cortex of these animals had normal single unit activity, but there was a striking absence of binocular cells (Fig. 11*a*).

Tangential penetrations showed that the monocular cells were grouped in the usual regular columnar pattern (Fig. 4*c*), suggesting that binocular cells had lost the input from the nondominant eye. The fourth monkey had low acuity in one eye, and fewer cortical cells were driven from that eye than from the normal one (Fig. 11*b*). When in five additional monkeys a strabismus was produced at later times during the critical period, there was an increase in the proportion of binocularly driven cells. From these results and earlier experiments in cats and monkeys, it seems that the period during which the cortex can be influenced by the artificially induced strabismus is comparable in duration and sensitivity to that observed with monocular deprivation⁴⁸⁻⁵⁰.

The binocular deprivation and strabismus experiments support the notion that competition, rather than disuse, is the main cause of the observed changes⁴³. The right circumstances must exist, however, for the competition to occur, since cells in the normal monkey tend to be dominated by one eye or the other¹⁰, and the dominant eye does not take over the cell completely. The difference between normal and deprived animals is that in normal conditions a cell receives input synchronously from the two eyes, whereas in monocularly deprived animals the two eyes do not act together. The maintenance of a given input may depend on the rate of firing of the postsynaptic cell while

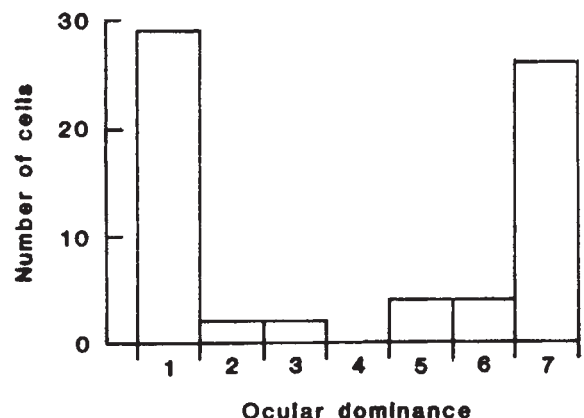


Fig. 10 Ocular dominance histogram of a monkey with binocular lid suture from birth to 30 days of age. The number of binocular cells was low (from ref. 44).

that input is active³ so that in normal animals the nondominant input is maintained by the activity of the dominant input. Carefully designed experiments by Singer *et al.*⁵¹ and Wilson *et al.*⁵² have provided support for the notion that it is crucial to activate the postsynaptic cell in order to change ocular dominance (for a more general discussion of synapse formation and stabilization see refs 53, 54).

In addition to providing insight into the mechanisms of development and plasticity in the visual cortex, the strabismus experiments may be of direct clinical relevance. A common situation in children with strabismus is that they have good vision in both eyes, but cannot fuse the images in the two eyes. These children often use the two eyes alternately, fixating and attending first with one eye and then with the other. The lack of binocular cells in strabismic animals is perhaps the physiological basis of this condition^{12,47}. Another common consequence of strabismus in children is a loss of acuity in one eye (strabismic amblyopia). The physiological mechanism of this condition is less well understood, even if our experiments (Fig. 11*b*) and those of others^{12,55} indicate that one eye has been weakened in its ability to drive cortical cells, as is seen in monocular deprivation.

As mentioned above, the late monocular deprivation in the monkey (Fig. 5*c*) and reversal experiments (Figs 8*b*, 9*c*) altered the cortical circuit at stages subsequent to the input from the LGN. Another series of experiments illustrated this point dramatically. The approach is a variation of the original experiments by Hirsch and Spinelli⁵⁶ and by Blakemore and Cooper⁵⁷ in which kittens were raised viewing only stripes of one orientation. In our experiments we allowed a monkey to see vertical stripes through one eye only (T.N.W., M. Carlson and D. H. Hubel, in preparation). The other eye was deprived by lid suture. This effectively produced a different condition of deprivation for different populations of cortical cells: those with vertical orientation preference were monocularly deprived, and those with horizontal orientation preference were binocularly deprived. We recorded from the striate cortex after 57 h of exposure (between days 12 and 54 after birth) and found normal levels of activity and cells of all orientations with their usual regular columnar arrangement¹⁷. There was an overall domin-

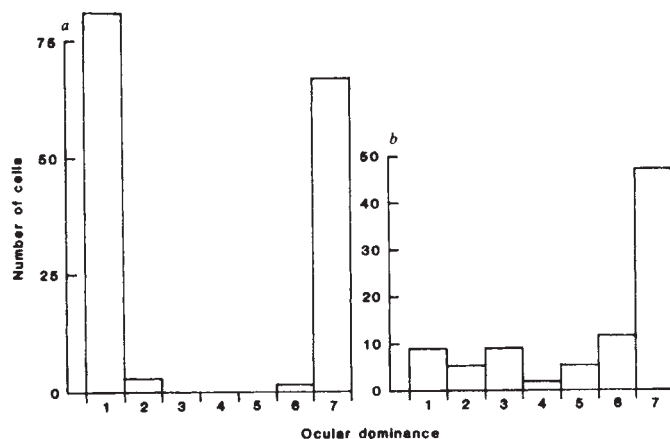


Fig. 11 Ocular dominance histograms of cells recorded in the striate cortex of two strabismic rhesus monkeys. *a*, Eye preference of cells recorded in a 3 yr-old monkey in which the lateral rectus of the right eye was sectioned at 3 weeks of age. Binocular cells are almost completely absent; the cells are driven exclusively either by the right or the left eye. Cells are clustered in a columnar fashion (Fig. 4*c*). The monkey had a 10° convergent strabismus and normal acuity in both eyes, but it could not fuse images presented separately to the two eyes. *b*, A monkey with convergent strabismus resulting from section of the lateral rectus muscle of the right eye at 3 weeks. Behavioural testing showed normal acuity in the left eye (20/30) and lower acuity in the right eye (20/60 to 20/120). There was no difference in refraction between the two eyes. The amblyopic eye influenced fewer neurones in the superficial and deep layers of the striate cortex. The ocular dominance columns in layer 4C had appeared normal when examined in tangential sections stained with a reduced silver method (Liesegang)¹⁹ (T.N.W. and D. H. Hubel, unpublished).

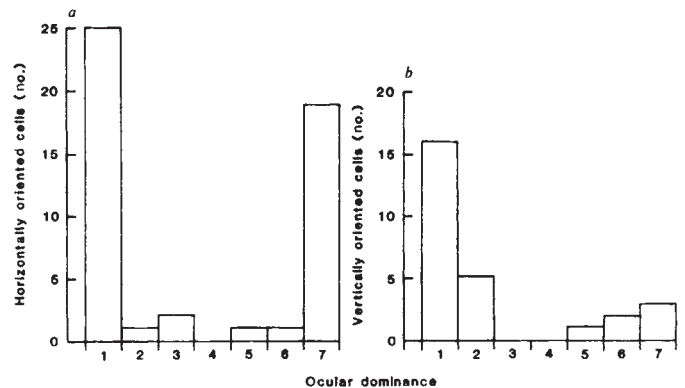


Fig. 12 Ocular dominance histograms of cells recorded from a monkey with the right eye closed at 12 days of age, then kept in the dark except for 57 h of self-exposure to vertical stripes during the subsequent 42 days. *a*, Cells ($n = 48$) recorded in the right striate cortex with preferred orientation within 45° of the horizontal axis. Few binocular cells and a good number of monocular cells responded to stimulation of either the left or the right eye. The distribution of eye preference resembled that seen in binocularly deprived animals (compare with Fig. 10). *b*, Cells ($n = 27$) with preferred orientation within 45° of the vertical axis. The majority of the recorded cells responded to the open eye, producing a histogram similar to that seen after monocular deprivation (T.N.W., M. Carlson and D. H. Hubel, in preparation).

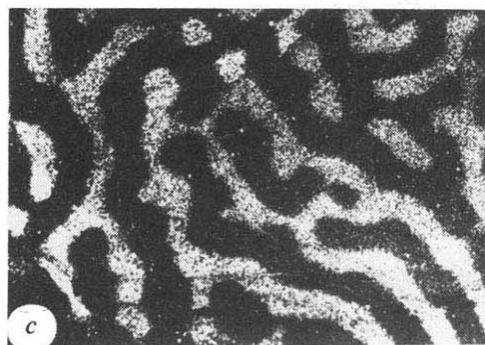
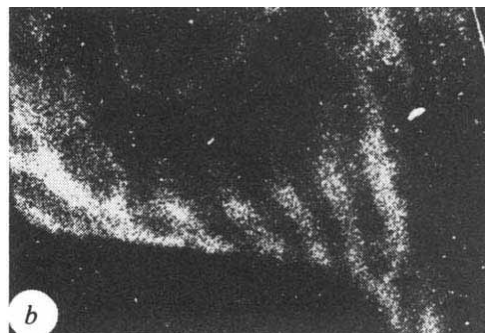
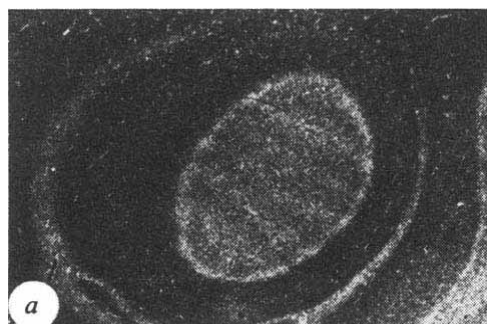
ance of the open eye, but when we produced separate ocular dominance histograms for vertically and horizontally oriented cells (Fig. 12), it became clear that horizontally oriented cells tended to be driven monocularly by either eye (a picture typical of binocular deprivation), and vertically oriented cells tended to be driven monocularly by the exposed eye only (a picture typical of monocular deprivation). Thus these findings again demonstrate that in addition to influencing the thalamocortical input, deprivation can alter the connections in the cortex without changing that input. This question has also been addressed in the kitten by rather different approaches, but with essentially the same results^{58,59}.

In looking at the effects of various forms of deprivation, one gets certain insights into the processes that govern the balance between different inputs, enabling the visual cortex to integrate information in the appropriate manner. We have learned that competition and synchronization of inputs are important factors in forming and maintaining this balance. If these processes are disturbed early in life, the system can be permanently altered.

Normal development

We cannot properly evaluate the experiments on visual deprivation without having detailed knowledge about the normal development of the visual system. To assess the relative importance of the genetic programme and the visual environment, it is necessary to evaluate the capabilities of the visual cortex at birth. The monkey is visually alert at or very soon after birth, and cells in the visual cortex show orientation preference and binocularity as in the adult monkey. This was shown by single-cell recordings in neonatal monkeys with no experience of contours of forms⁴⁴.

Compared with monkeys, kittens are poorly developed at birth; their eyes do not open until the second week, and they spend their first 3 weeks mainly eating and sleeping. In the visual cortex, cells tend to give weak or erratic responses during the early postnatal period and come to respond like adult cells at about 3–4 weeks of age^{9,29,60}. During the same time period, cortical cells differentiate and active synapse formation takes place^{61,62}. Whether cells in the cat visual cortex develop normally without visual experience, as was originally reported for binocularly sutured kittens⁸ was questioned at first⁶³, but subsequently confirmed in several studies^{60,64–66}. Whether in the kitten all cortical cells can develop fully through innate mechanisms is not entirely clear, as animals raised in the dark or with binocular lid closures seem to have a certain fraction of unresponsive or unoriented cortical cells^{60,64,66}.



1 mm

Fig. 13 Dark-field autoradiographs of geniculate afferent terminals in the striate cortex of normal neonatal monkeys in which the right eye had been injected with a radioactive tracer 1 to 2 weeks earlier. *a*, Single section from the left hemisphere of a normal 6-day-old monkey; the section grazes layer 4C tangentially in the central oval region. Silver grains are distributed continuously over the layer 4C, bands of alternating higher and lower grain density indicate that afferents for the two eyes are in the process of columnar segregation. *b*, A single section through layer 4C of the left hemisphere of a normal 3-week-old monkey. The slight blurring of the margin of the labelled bands suggests a modest intermixing of left- and right-eye afferents at the borders of ocular dominance columns. *c*, Autoradiographic montage of the geniculate labelling pattern in layer 4C of the right striate cortex of a 6-week-old normal monkey. The ocular dominance columns appear as sharply defined as they do in the adult monkey (adapted from ref. 15).

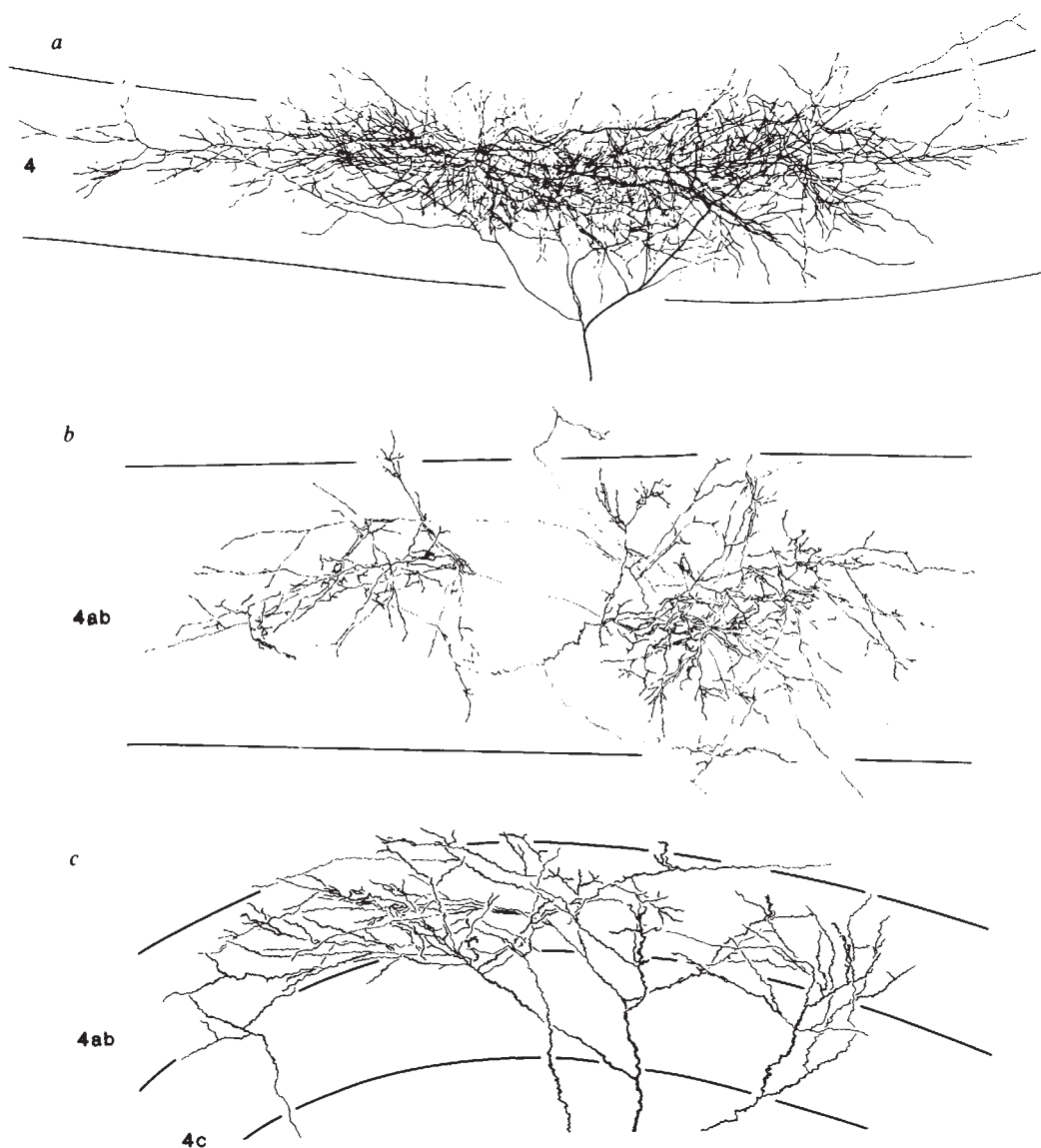


Fig. 14 Patterns of arborization of single geniculate axons in layer 4 of the cat visual cortex. *a*, A 17-day-old kitten. The arborization of a single afferent is shown at an age prior to columnar segregation⁶⁹. The axon, impregnated in its entirety with the rapid Golgi method, arborizes profusely and uniformly over a disk-shaped area more than 2 mm in diameter (from Ref. 71). *b*, Normal adult cat. OFF-centre geniculate afferent (Y-type)⁷⁵ injected intra-axonally with horseradish peroxidase in the striate cortex. The arborization, which is entirely within layer 4 AB, forms two patches separated by a terminal-free gap. Presumably this pattern corresponds to the segregation of the input from the two eyes in a columnar fashion (from ref. 73). *c*, A monocularly deprived cat between 2 weeks and more than 1 yr. The cell is a nondeprived Y-type afferent with an ON-centre receptive field injected intra-axonally with horseradish peroxidase. The arborization is primarily within layer 4AB, but it does not have the normal patchy distribution of terminals. The absence of a terminal-free region indicates that nondeprived geniculate branches are present in the territory that normally belongs only to the other eye (unpublished observations).

The newborn animal does differ from the adult in one significant respect, relating to the segregation of the afferents from the two eyes in layer 4C. In the newborn monkey we were able to show by injection of ^3H -proline into the eye that the inputs from the two eyes strongly overlap with only a mild fluctuation of eye dominance in a bandlike pattern¹³. Des Rosiers *et al.*⁶⁷ confirmed this observation with the 2-deoxy-D-glucose method. In the monkey fetus, Rakic showed that initially the left- and right-eye afferents overlap completely, and not until a few days before birth do they begin to sort out into ocular dominance columns⁶⁸. We followed this process of segregation postnatally; it was completed by 4–6 weeks of age (Fig. 13)¹³. Recordings also indicated an initial overlap, followed by separation of the inputs from the two eyes in layer 4C, and the time course of the events was similar. The process of segregation did not require visual experience, since it also occurred in an animal raised in the dark¹⁵. In kittens, ocular dominance columns are formed much as they are in monkeys, showing a sequence of initial overlap and segregation during the first few months of life⁶⁹, even though in this species visual experience appears necessary for their normal development⁷⁰.

In the kitten, it has been possible to examine the segregation of ocular dominance columns at the level of single cells. In the early postnatal period, a single geniculate afferent gives off numerous branches innervating without interruption an area covering several future ocular dominance columns (Fig. 14a)⁷¹. As the axon matures, there appears to be a selective loss of branches, so that ultimately it innervates ocular dominance columns serving one eye and leaves gaps for the columns serving the opposite eye (Fig. 14b)^{72,73}. In a cat monocularly deprived during its critical period, a geniculocortical afferent with input from the normal eye is shown in Fig. 14c (unpublished observations). It appears to have innervated an area that normally would have been occupied by the other eye.

Both the autoradiography (Fig. 3) and the single-cell reconstructions (Fig. 14) suggest a mechanism for the expansion and contraction of ocular dominance columns in monocular closure. The terminals with input from the normal eye continue to occupy the territory which they would normally have relinquished, while the deprived terminals are trimmed to an abnormal extent. Another mechanism appears to operate at slightly later stages: when the ocular dominance columns are fully segregated (at 6 weeks of age in the monkey), monocular closure

still causes an expansion and a contraction of the columns similar to that seen after earlier closure (Fig. 6b). This argues for a mechanism of sprouting by one axon into the territory originally occupied exclusively by the deprived eye. Perhaps sprouting occurs from the axon branches that traverse the ocular dominance columns for the other eye (Fig. 14b). Reverse suture experiments also indicate that both trimming and sprouting are involved in plastic changes of the geniculocortical pathway (Fig. 9a–c). We know little of the biochemical mechanisms underlying these changes, except for the intriguing observation of the possible role of noradrenaline in neural plasticity⁷⁴. Since the critical period seems to vary in onset and duration between different brain regions and even between layers of an individual cortical area (compare layers 4C α and 4C β in monkey striate cortex¹⁵ (Fig. 9a, b), the control of plasticity appears to be specific and localized, not a phenomenon controlled by diffuse processes.

Conclusion

Innate mechanisms endow the visual system with highly specific connections, but visual experience early in life is necessary for their maintenance and full development. Deprivation experiments demonstrate that neural connections can be modulated by environmental influences during a critical period of postnatal development. We have studied this process in detail in one set of functional properties of the nervous system, but it may well be that other aspects of brain function, such as language, complex perceptual tasks, learning, memory and personality, have different programmes of development. Such sensitivity of the nervous system to the effects of experience may represent the fundamental mechanism by which the organism adapts to its environment during the period of growth and development.

I thank the staff, faculty and students in the Department of Neurobiology. For more than two decades this has been a unique place because of its blend of scientific excellence and compassion. The late Stephen Kuffler played a very special role in the creation of this environment. I thank both the Johns Hopkins and Harvard Medical Schools for providing me with invaluable opportunities for scientific training and research. I want also to recognize Robert Winthrop who provided the resources for my endowed professorship, and acknowledge the steady and generous support from the National Eye Institute. I thank A. Houdek and C. Gilbert for their help in preparing the manuscript.

- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106 (1962).
- von Senden, M. *Space and Sight* (The Free Press, Glencoe, 1960).
- Hebb, D. O. *The Organization of Behavior* (Wiley, New York, 1949).
- McCleary, R. A. *Genetic and Experimental Factors in Perception* (Scott Foresman, Glenview, 1970).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **148**, 574 (1959).
- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 1003 (1963).
- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 978 (1963).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **26**, 994 (1963).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **206**, 419 (1970).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215 (1968).
- Wiesel, T. N. & Hubel, D. H. *Symp. Int. Un. of physiol. Sci. Abstr. P.*, Munich (1971).
- Baker, F. H., Grigg, P. & von Noorden, G. K. *Brain Res.* **66**, 185 (1974).
- Hubel, D. H., Wiesel, T. N. & LeVay, S. *Phil. Trans. Soc. Lond.* **B278**, 377 (1977).
- Blakemore, C., Garey, L. J. & Vital-Durand, F. *J. Physiol., Lond.* **283**, 223 (1978).
- LeVay, S., Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **191**, 1 (1980).
- Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **146**, 421 (1972).
- Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **158**, 267 (1974).
- Wiesel, T. N., Hubel, D. H. & Lam, D. *Brain Res.* **79**, 273 (1974).
- LeVay, S., Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **159**, 559 (1975).
- Hubel, D. H. *Nature* **299**, 515–524 (1982).
- Hubel, D. H. & Wiesel, T. N. *Proc. R. Soc. B* **198**, 1 (1977).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28**, 229 (1965).
- Barlow, H. B. *Nature* **258**, 199 (1975).
- Pettigrew, J. D. in *Neuronal Plasticity* (ed. Cotman, C.) 311–330 (Raven, New York, 1978).
- Movshon, J. A. & Slyter, R. C. *Rev. Psychol.* **32**, 477 (1981).
- Wiesel, T. N. & Raviola, E. *Nature* **266**, 66 (1977).
- Guillery, R. W. & Stelzner, D. J. *J. comp. Neurol.* **139**, 413 (1970).
- Crawford, M. L. J., Blake, R., Cool, S. J. & von Noorden, G. K. *Brain Res.* **84**, 150 (1975).
- Blakemore, C. & von Slyter, R. C. *J. Physiol., Lond.* **237**, 195 (1974).
- Arden, G. B. & Barnard, W. M. *Trans. ophthalm. Soc. U.K.* **99**, 419 (1979).
- Vaegan, D. T. *Trans. ophthalm. Soc. U.K.* **99**, 432 (1979).
- von Noorden, G. K. *Am. J. ophthalm.* **92**, 416 (1981).
- Dews, P. B. & Wiesel, T. N. *J. Physiol., Lond.* **206**, 437 (1970).
- Ganz, L., Hirsch, H. V. B. & Tieman, S. B. *Brain Res.* **44**, 547 (1972).
- Mitchell, D. E., Cynader, M. & Movshon, J. A. *J. comp. Neurol.* **176**, 53 (1977).
- Olson, C. R. & Freeman, R. D. *J. Neurophysiol.* **41**, 65 (1978).
- Awaya, S., Sugawara, M. & Miyake, S. *Trans. ophthalm. Soc. U.K.* **99**, 447 (1979).
- Odom, J. V., Hoyt, C. S. & Marg, E. *Archs Ophthalm.* **99**, 1412 (1981).
- Beller, R., Hoyt, S., Marg, E. & Odom, J. V. *Am. J. Ophthalm.* **91**, 559 (1981).
- Jacobson, S. G., Mohindra, J. & Held, R. *Br. J. Ophthalm.* **65**, 727 (1981).
- Blakemore, C., Vital-Durand, F. & Garey, J. *Proc. R. Soc. B* **213**, 399 (1981).
- Swindale, N. V., Vital-Durand, F. & Blakemore, C. *Proc. R. Soc. B* **213**, 435 (1981).
- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **28**, 1029 (1965).
- Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **158**, 807 (1974).
- Guillery, R. W. *J. comp. Neurol.* **144**, 117 (1972).
- Guillery, R. W. *J. comp. Neurol.* **148**, 417 (1973).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28**, 1041 (1965).
- Yinon, U. *Exp. Brain Res.* **26**, 151 (1976).
- Crawford, M. L. J. & von Noorden, G. K. *Invest. Ophthalm.* **18**, 496 (1979).
- Jacobson, S. G. & Ikeda, H. *Exp. Brain Res.* **34**, 11 (1979).
- Singer, W., Rauschecker, J. & Werth, R. *Brain Res.* **134**, 568 (1977).
- Wilson, J. R., Webb, S. V. & Sherman, S. M. *Brain Res.* **136**, 277 (1977).
- Stent, G. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 997 (1973).
- Changeux, J.-P. & Danchin, A. *Nature* **264**, 705 (1976).
- Ikeda, H. & Wright, M. J. *Exp. Brain Res.* **25**, 63 (1976).
- Hirsch, H. V. B. & Spinelli, D. N. *Science*, **168**, 869 (1970).
- Blakemore, C. & Cooper, G. F. *Nature* **228**, 477 (1970).
- Cynader, M. & Mitchell, D. E. *Nature* **270**, 177 (1977).
- Rauschecker, J. P. & Singer, W. *Nature* **280**, 58 (1979).
- Buisseret, P. & Imbert, M. J. *Physiol., Lond.* **255**, 511 (1976).
- Marty, R. *Archs Anat. Microsc. Morphol. Exp.* **52**, 129 (1962).
- Cragg, B. G. *J. comp. Neurol.* **160**, 147 (1975).
- Pettigrew, J. D. *J. Physiol., Lond.* **237**, 49 (1974).
- Blakemore, C. & von Slyter, R. C. *J. Physiol., Lond.* **248**, 663 (1975).
- Sherk, H. & Stryker, M. P. *J. Neurophysiol.* **39**, 63 (1976).
- Singer, W. & Tretter, F. J. *Neurophysiol.* **39**, 613 (1976).
- Des Rosiers, M. H. *et al. Science* **200**, 447 (1978).
- Rakic, P. *Phil. Trans. R. Soc. B* **278**, 245 (1977).
- LeVay, S., Stryker, M. P. & Shatz, C. J. *J. comp. Neurol.* **179**, 223 (1978).
- Swindale, N. V. *Nature* **290**, 332 (1981).
- LeVay, S. & Stryker, M. P. *Soc. Neurosci. Symp.* **4**, 83 (1979).
- Ferster, D. & LeVay, S. *J. comp. Neurol.* **182**, 923 (1978).
- Gilbert, C. & Wiesel, T. N. *Nature* **280**, 120 (1979).
- Kasamatsu, T. & Pettigrew, J. D. *J. comp. Neurol.* **185**, 139 (1979).
- Enroth-Cugell, C. & Robson, J. G. *J. Physiol., Lond.* **182**, 923 (1966).

REVIEW ARTICLE

Tapping the immunological repertoire to produce antibodies of predetermined specificity

Richard A. Lerner

Committee for the Study of Molecular Genetics, Research Institute of Scripps Clinic, La Jolla, California 92037, USA

We understand the structure of antibodies in detail, but know little about the molecular basis of the immunogenicity of proteins. Recent experiments have shown that chemically synthesized peptides representative of virtually any part of the surface of a protein can elicit antibodies reactive with the native molecule. Such peptides can serve as synthetic vaccines, and antibodies, useful in the study of changes in protein structure, can be generated. As these antibodies react with regions of the protein known in advance to the experimenter, they can be said to be of predetermined specificity.

THE immune system of a mammal is one of the most versatile systems in the biological kingdom as probably greater than 1.0×10^7 antibody specificities can be produced¹. Indeed, much of contemporary biological and medical research is directed toward tapping this repertoire. As it is so vast, it might appear to be a relatively simple matter to produce antibodies of a particular specificity but until recently this was not the case because of two essential complications. The first problem is that serum antibodies consist of a mixture of molecules of diverse reactivity. As such they are useful for studying whole proteins but an understanding of fine specificity is difficult if not impossible. The development by Köhler and Milstein of the hybridoma methodology has solved this problem by making it possible to obtain in pure form antibodies of a single specificity from those induced during an immune response² but this left a second problem in that during an immune response to an intact protein antibodies are only produced against a very limited set of determinants within the protein molecule. This problem is the one to be discussed here.

Limited response to intact proteins

The key to eliciting antibodies of predetermined specificity is an understanding of what constitutes an immunogenic determinant on a protein. Whereas the term an antigenic determinant simply reflects the ability of a region in a protein to bind to antibodies, an immunogenic determinant is a region capable of inducing antibody. (Here we further define an immunogenic determinant as one that induces antibody reactive with the native protein.)

Previous studies on the nature of antigenic determinants, mostly following immunization with intact proteins, have led to two fundamental conclusions (reviewed in ref. 3, see also refs 4–21). The first is that during an immune response to a native protein, antibody reactivity is confined to only a few regions of the molecule. Studies on enzymatically fragmented proteins suggested that most globular proteins contain fewer than five antigenic sites; as a rough rule, about one site per 5,000 daltons of protein. The second conclusion is that antigenic determinants are dependent on tertiary conformation and are often constructed from discontinuous regions of the protein chain brought into proximity by folding of the molecule. These determinants are called 'conformational' or 'discontinuous'. Thus, by the mid-1970s, a picture of the antigenic profile of a protein had emerged. The determinants are few in number and largely dependent on native conformation (Fig. 1). However, throughout these studies it was tacitly assumed that antigenicity and immunogenicity are equivalent; in other words, the number of antigenic determinants of an intact protein was presumed to set a limit on the number of protein fragments which would carry immunogenic determinants.

The repertoire can be tapped

Needless to say, the above concepts did not bode well for a general method of producing antibodies reactive with most regions of a protein molecule. Our own interest in the problem followed an experiment we carried out to detect a protein potentially encoded by the Moloney leukaemia virus genome. During our sequencing of this genome, we found an open reading frame whose coding capacity did not fit with the known biochemistry of the viral proteins, a problem we have called genotype in search of phenotype. We decided to synthesize chemically a peptide from within the protein predicted by the nucleotide sequence, raise an antibody to it, and see if that antibody reacted with protein(s) in infected cells. Because of uncertainties as to how or if RNA splicing might be taking place, we synthesized a peptide potentially encoded by the 3' end of the reading frame, and thus representative of the C-terminus of the putative protein. Indeed, the anti-peptide antibody precipitated a protein from infected but not normal cells²². Such an approach could be very useful as more and more DNA sequences were generated, but it was far from proven as a general methodology. We had studied the C-terminus of a protein and it seemed possible that the method might be useful only in detecting the ends of proteins. It was thought that the untethered C-terminus of a protein was relatively free to rotate and could be thought of as a kind of hapten carried by the rest of the molecule, a situation which we could have fortuitously duplicated when we coupled the peptide to the carrier protein for immunization.

To test the generality of the method we carried out a study on a protein of known structure. We chose the influenza virus haemagglutinin (HA) because the complete nucleic acid sequence of its gene is available²³ and its crystallographic structure is known at high resolution²⁴. A series of peptides covering 75% of the HA1 chain were chemically synthesized (we have now synthesized additional peptides so that the entire span of the molecule is represented) and antibodies made to each of the peptides. Antibodies to almost all (18 of 20) peptides react with the native molecule²⁵. Because in its folded state the HA1 molecule displays a number of conformations including α -helix, random coil and β -sheets, it is clear that the ability of the anti-peptide antibodies to react with the intact structure is independent of any particular conformation or location in the molecule²⁵. Probably the only requirement for selecting an immunogenic peptide is that a part of the sequence be located on the surface of the molecule so as to be available to antibody. In Fig. 2b, we show the portion of the surface of the HA1 molecule against which we have made antibodies using chemically synthesized peptides as immunogens. In Fig. 2a, we show those areas of the molecule thought to be immunogenic during viral infection or immunization with intact virus or purified viral

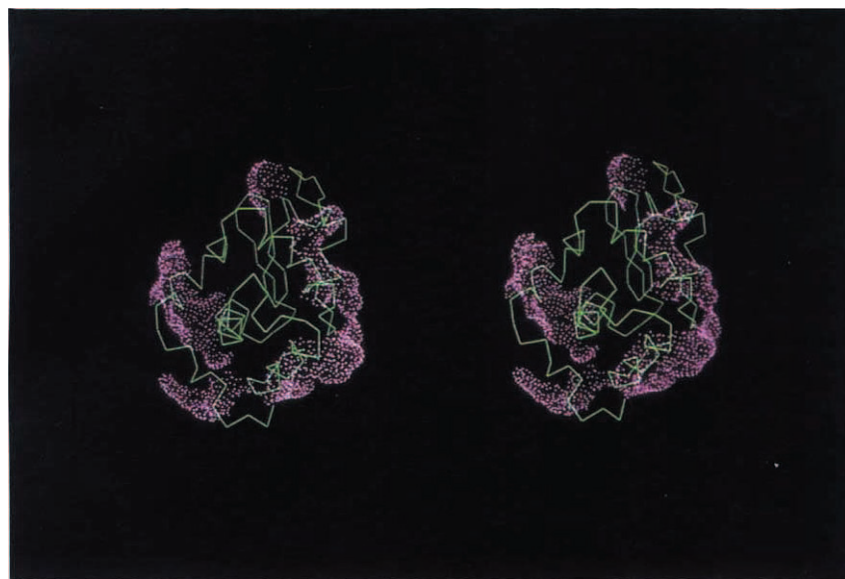


Fig. 1 Exposed surfaces of antigenic sites (shown in pink) of hen's egg-white lysozyme. Stereo projections showing the limited number of antigenic sites in the lysozyme molecule and their discontinuous ('conformational') nature. The data are based on Atassi and Lee⁷¹ using the structure of Blake *et al.*⁷² and Imoto *et al.*⁷³. Exposed surfaces based on α -carbon positions using molecular surface computer program of Connolly⁷⁴. The three antigenic sites are constructed by the spatially contiguous residues as follows: (1) Arg 125, Arg 5, Glu 7, Arg 14 and Lys 13; (2) Trp 62, Lys 97, Lys 96, Asn 93, Thr 89 and Asp 87; (3) Lys 116, Asn 113, Arg 114, Phe 34 and Lys 33.

a

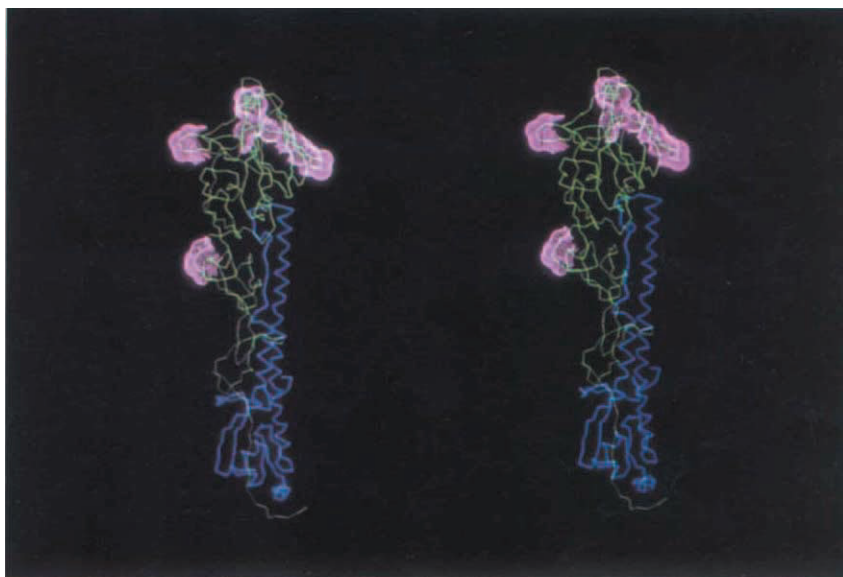
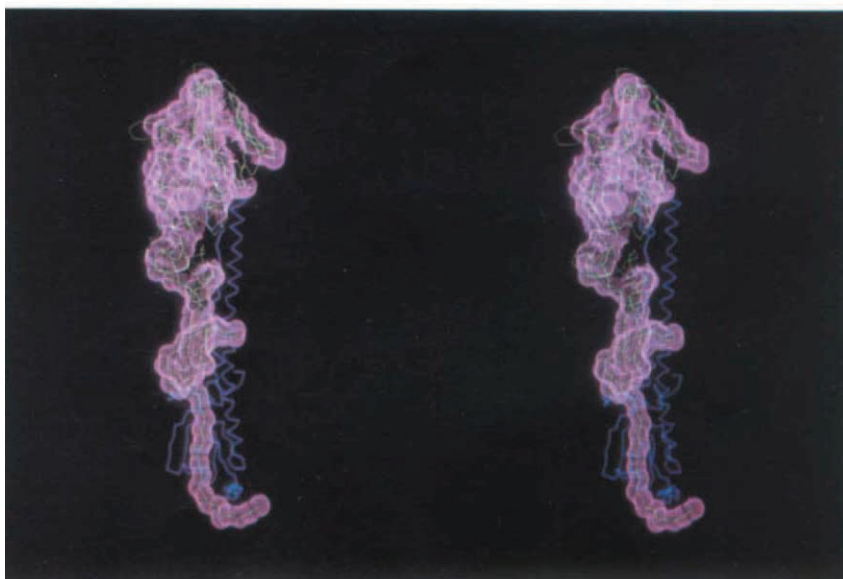


Fig. 2 Antigenic and immunogenic sites (shown in pink) of the influenza virus HA1 molecule. (HA1 shown in green, HA2 shown in blue.) Structure of the molecule based on crystallographic coordinates of Wilson *et al.*³¹. Exposed surfaces based on α -carbon positions using molecular surface computer program of Connolly⁷⁴. *a*, Stereo-pair representing sites eliciting antibodies during the process of infection or immunization with intact virus. *b*, Stereo-pair representing sites against which antibody can be induced by immunization with chemically synthesized peptides.

b



proteins²⁶⁻²⁸. The conclusions from these studies were clear—the immunogenicity of a protein is less than the sum of the immunogenicity of its pieces²⁵ with, however, one caveat. The map in Fig. 2 is based on neutralization data and it is possible that during ordinary immunization, antibodies reactive with other parts of the protein are generated but are not scored because they are not neutralizing. (This problem arises because not all antibodies that bind to viral proteins inhibit the infectivity of the virus.) If this were the case then the collection of anti-virus antibodies would have a reactivity pattern much broader than that observed by neutralization studies. This is, however, probably not the case: we have found that high-titred antibody made against the intact haemagglutinin does not react with any of the synthetic peptides²⁵. We recently carried out a more compelling demonstration of the exclusion of some reactivities in anti-virus sera by taking advantage of the facts that among various influenza strains there are constant and variable regions of the HA1 and HA2 components of the viral haemagglutinin, and that antibody to the native molecule does not widely neutralize across strains. If an anti-peptide antibody to a conserved region neutralized across strains whereas an anti-virus antibody did not, it would indicate that different immunological specificities are generated during the two types of immunization. We therefore studied antibodies to several peptides from conserved portions of the protein structure and showed that even though the anti-virus antibody has a titre against the homologous strain which is about 100-fold higher than that of the anti-peptide antibodies, only the latter neutralizes across strains (S. Alexander and R.A.L., unpublished). Thus, even during a vigorous immune response against virus, the region represented by these synthetic peptides is not immunogenic; hence, by using peptide immunization one can generate antibody specificities that cannot be obtained in any other way.

Antibody of predetermined specificity in biology

The spread of the use of chemically synthesized peptides to generate antibodies of predetermined specificity is indicated by the number and diversity of experiments recently carried out. The antibodies have been designed for a wide variety of uses.

Detection of proteins predicted from nucleic acid sequences: Anti-peptide antibodies have proved useful in detecting proteins predicted on the basis of nucleic acid sequences to be present in cells. The technology has been particularly successful in discovering DNA and RNA tumour virus proteins implicated in cell transformation. Anti-peptide antibodies have been used to detect the large T antigen of polyoma and SV40 viruses, as well as the cellular transformation-associated protein uniquely expressed in many types of malignant cells²⁹⁻³¹. Green and Brackmann (personal communication) made an anti-peptide antibody that precipitates the 53,000 molecular weight (53K) protein encoded by the adenovirus E1B transcription unit. As expression of the 53K protein is essential for a fully transformed cell, this antibody together with that made to adenovirus E1A products (see below) should be useful in studying the process of cell transformation. A particularly successful use of anti-peptide antibodies has been in defining the proteins encoded by the oncogenes of the rapidly transforming retroviruses including those of the Moloney sarcoma^{32,33}, feline sarcoma³⁴, Rous sarcoma (refs 35, 36 and L. E. Gentry *et al.*, personal communication), avian myeloblastosis³⁷ and Simian sarcoma virus³⁸.

Antibodies against functionally active regions of proteins: Peptide hormones are often cleaved from larger precursor proteins, which contain multiple hormones. Anti-peptide antibodies have been used to localize the portions of the 31,000-MW γ -melanocyte-stimulating hormone and 17,500-MW calcitonin precursor which correspond to the functional activities of these two hormones^{39,40}.

Schaffhauser *et al.* have used anti-peptide antibodies to perturb the functional activity of proteins⁴¹. The middle T antigen of polyoma virus has been implicated in cell transformation

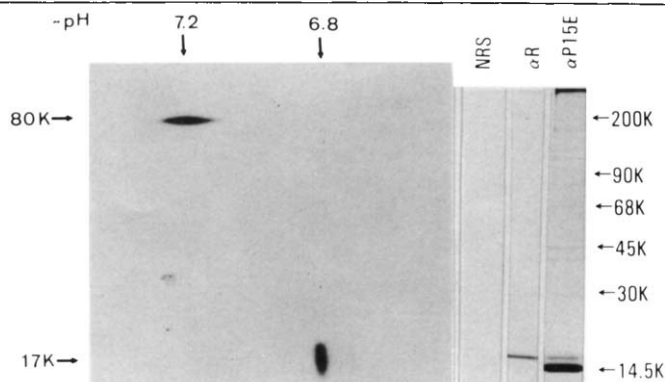


Fig. 3 Studies illustrating the use of anti-peptide antibodies to demonstrate the cleavage of viral protein during virus maturation. Left: immune precipitates of virus infected cells with anti-peptide antibody. Extracts of ³H-leucine-labelled cells were reacted with anti-peptide antibody and the solubilized immune precipitates characterized by two-dimensional gel electrophoresis. The 80K protein is the entire product (Pr80^{env}) of the envelope gene and contains N-gp70-15E-C. The 17K protein (pre-15E) was confirmed by radioactive protein sequencing to be the precursor to the 15E anchor protein⁴⁴. Right: immune precipitates of labelled virus showing that amino acids had been cleaved from the C-terminus of pre-15E during virus maturation to yield viral p15E. Virus labelled with ³⁵S-methionine was reacted with normal rabbit sera (NRS); anti-peptide antibody [here called α R to designate that this serum represents antibody against amino acids predicted by right most (3') possible 15 codons of the viral genome]; and antibody against the mature anchor protein p15E (α P15E). The bulk of the p15E in virus fails to react with α R sera and thus lacks a portion of its C-terminus⁴⁴.

and contains a tyrosine kinase activity that predominantly phosphorylates tyrosine 315 of the middle T protein. They synthesized a nonapeptide corresponding to residues 311–319 of the middle T antigen. This antibody not only precipitates middle T in infected cells but inhibits phosphorylation of the protein *in vitro*. Interestingly, the anti-peptide antibody does not react with the middle T antigen when tyrosine 315 is phosphorylated. Shinnick *et al.*⁴² have also used anti-peptide antibodies to localize and perturb the functions of enzymes encoded by viral genomes. The retroviral *pol* gene is about twice as big as necessary to encode the 80,000-MW reverse transcriptase⁴³. Gene products were sampled by synthesizing peptides at intervals of about 100 amino acids along the predicted *pol* sequence and antibodies against these peptides used to study *pol* gene products in infected cells. Antibodies to peptides predicted from the 5' end of the gene precipitated an 80,000-MW protein and inhibited the reverse transcriptase activity whereas antibodies to peptides from the middle of the gene precipitated a 40,000-MW protein and inhibited the virus-associated endonuclease activity. Antibodies to peptides from the 3' end of the gene precipitate a 20,000-MW protein which according to enzyme inhibition studies seems to be protease. Thus, anti-peptide antibodies are useful not only in detecting the protein product of a gene, but also in identifying its function.

Following protein domains: Anti-peptide antibodies have been used in the difficult problem of following the fate of individual protein domains and have shown that the cleavage and removal of the hydrophilic C-terminus from the retroviral membrane anchor protein takes place during virus maturation⁴⁴. The main envelope glycoprotein, gp70, is anchored into the viral membrane through its attachment by disulphide bonds to the membrane spanning proteins, p15E. Antibody against a synthetic peptide corresponding to the C-terminus of the pre-15E protein was made and used to study the fate of this region of the protein during processing leading to virus formation. In infected cells the antibody detected two proteins of molecular sizes of 80K and 17K corresponding, respectively, to the envelope polypeptide precursor containing gp70 and p15E, and one of the cleavage products (pre-15E) (Fig. 3). However, when the proteins of radioactively labelled virus were studied

with the anti-peptide antibody, it could be seen that the C-terminus predicted from the nucleotide sequence was missing from the mature p15E protein (Fig. 3), thus demonstrating its removal during virus maturation. It remains to be seen if cleavage and processing of C-termini are frequent events which will be found in other systems or if they are peculiar to events involved in viral maturation.

Baron and Baltimore⁴⁵ and Semler *et al.*⁴⁶ used anti-peptide antibodies to follow poliovirus protein domains, by making antibodies to chemically synthesized peptides corresponding to the genome-linked protein (VPg). Since mature poliovirus proteins are generated by a cascade of cleavages starting with a large precursor (NCVPOO), the peptide of interest is located in various positions in the different intermediate cleavage products, ultimately becoming the C-terminus of a 12,000-MW polypeptide from which 22 amino acids are donated to the 5' end of the genomic RNA. In the Baron and Baltimore experiments antibodies raised against the entire 22 amino acids of VPg as well as a peptide corresponding to its 14 most C-terminal amino acids reacted with the peptide when it was present in five different members of the cleavage cascade. As well as providing the solution of a biological problem, these results demonstrate that the reaction of anti-peptide sera with the native protein is relatively independent of the position of the target in the native structure.

Exon usage: A special example of the use of antibodies of predetermined specificity to follow protein domains is in the study of exon usage during gene expression. Shinnick and Blattner⁴⁷ used anti-peptide antibodies to follow exon usage in the immunoglobulin-D system. They synthesized peptides uniquely corresponding to the protein encoded by the exons specific for either the secreted or the membrane-bound form of IgD. Each of the anti-peptide antibodies was shown to be specific for one of the alternative forms of IgD and can now be used to follow the fate of these two proteins in cells. These results, again, demonstrate the production of specific reagents which would be difficult to make by other means. Anti-peptide antibodies have also facilitated the study of exon usage in the adenovirus-2E1A transcription unit. E1A encodes functions which both regulate expression of other early viral genes and have a role in cell transformation⁴⁸⁻⁵¹. The E1A region transcript is processed into at least two overlapping mRNAs (12S and 13S) which share 5' and 3' termini and differ by 138 nucleotides. Since the 12S and 13S mRNAs are in the same reading frame and translation is probably initiated at the common first AUG, it has been assumed that the two proteins encoded by these messages have common N- and C-termini and differ by 46 amino acids unique to the middle of the larger protein⁵². Feldman and Nevins prepared antibody to a 13 amino acid long synthetic peptide predicted from the nucleotide sequence to correspond to a hydrophilic portion of the putative 46 amino acids unique to the larger protein and were able to show that this antibody only reacted with the larger of the two proteins encoded in E1A⁵³. This antibody should be very helpful in defining the role of the different E1A proteins in transformation and control of transcription.

Anti-'wrong' reading frame antibodies: We prepared anti-'wrong' reading frame antibodies to study frameshift mutations in viral proteins (A. Sen and R.A.L., unpublished). We used the nucleotide sequence of several sarcoma virus transforming genes to predict the protein sequence which would correspond to +1 and +2 frameshifts far enough to the 3' end of the genome so that premature terminations would not occur. Using these antibodies we were able to detect frameshifted proteins in infected cells. Coupled with transfection of genes, such anti-'wrong' reading frame antibodies may be very useful in comparing the fate of wild-type and mutant proteins in the same cells.

Antibody of predetermined specificity in medicine

Small molecules were first used in the design of synthetic vaccines in 1938 when Goebel made antibody to the carbo-

hydrate antigen diazotized *p*-aminobenzyl cellobiuronide coupled to horse serum globulin^{54,55}. Mice immunized with the preparation acquired active resistance to infection with virulent type III pneumococci⁵⁴. Anderer, working from the results of denaturation and cleavage experiments and assuming that the immunogenicity of the tobacco mosaic virus protein depends on retention of conformation, studied a C-terminal hexapeptide coupled to bovine serum albumin. He showed that antibody to the hexapeptide would precipitate and neutralize the virus⁵⁶. Working with phage, Langbeheim and his colleagues synthesized two peptides from the coat protein of MS-2 and made rabbit antiserum against them⁵⁷. They were able to show binding by one of the anti-peptide antibodies. In their experiment, addition of the rabbit anti-peptide antibody was followed by addition of antisera against rabbit immunoglobulin, thus obscuring whether the anti-peptide antibody had any primary neutralizing activity.

Before synthetic vaccines could be realistically designed to combat eukaryotic viruses, two things were required: a method for obtaining protein sequences of relatively scarce viral proteins, and constraints of the actual or perceived need for conformational determinants had to be overcome. Advances by Sanger and Gilbert and their colleagues in nucleic acid sequencing technology solved the problem of obtaining reliable amino acid sequences of viral proteins, and chemical synthesis of peptides allowed production of antibodies not necessarily restricted, to reactivity with conformational determinants within native proteins. Recently, peptides corresponding to influenza HA1 (refs 25, 58), the hepatitis B surface antigen (HB_sAg)⁵⁹⁻⁶², the foot-and-mouth disease (FMDV) VPI protein⁶³, and the rabies virus glycoprotein⁶⁴, have been synthesized. One interesting result in the HB_sAg system is that peptides differing in only two residues were capable of inducing antibodies in rabbits and chimpanzees which were capable of distinguishing the y and d subtypes of HB_sAg⁶⁵. It appears that synthetic peptides can duplicate serological markers known to be of significance from classical virological and epidemiological studies. In the influenza⁶⁶ and FMDV systems⁶³, synthetic peptides induced neutralizing antibodies. Further studies using antibodies of predetermined specificity should, in turn, improve our understanding of the process of virus neutralization and thus guide the design of useful peptides. One can envisage that the reason for strain variation in viral proteins is due to selective pressures of the immune system, and thus that the variable regions signal areas available for, and sensitive to, antibody binding. Alternatively, one can aim for invariant regions with functional activities. An additional theoretical issue important to the design of new vaccines is that an intact protein when free may fold differently from when it is part of a virus particle and thus may not confront the immune system in such a way as to induce neutralizing antibodies. In these cases, synthetic peptides may offer advantages over purified viral capsid proteins (so-called subunit vaccines) because they can induce anti-virus antibodies which are independent of protein folding and can be directed to neutralizing sites on the virus surface.

Practical considerations

Given a nucleotide sequence, the problem often is which peptide to synthesize. As so many different peptides have been shown to induce antibodies reactive with the native protein, the only essential point is to choose one which contains hydrophilic amino acids and is thus likely to be exposed on the surface of the intact molecule. Also, peptides containing hydrophilic groups are likely to be soluble, making their handling and coupling to carrier molecules easier. A number of computer programs have been used to predict secondary structures⁶⁷⁻⁶⁹, which indicate regions of proteins located on the surface of the molecule and available to antibody but a simple search of the amino acid sequence for charged residues seems to be equally effective for most studies. In the experiments carried out so far, several different carriers and coupling methods have been used with equally good results, and so this choice does not seem

to be critical as long as an immunogenic carrier is used. However, where cultured cells are used, bovine serum albumin should be avoided as a carrier since it is a component of the growth media to which many cellular proteins bind non-specifically and the presence of antibodies to it can confuse immunoprecipitation studies of radioactively labelled cells.

Another question is that of the titre of an anti-peptide antibody compared with that raised by immunization with the intact protein. The answer, as far as we know, is that sometimes it is greater and sometimes less. However an important point is that by using peptide immunization, one can generate antibody specificities which cannot be obtained in any other way.

The minimum size of the peptide chosen is important and should be larger than six amino acids. We generally synthesize peptides of 15 amino acids. Considerably larger peptides have also proved useful²⁵ but often do not offer any advantage as one risks the problem that they are more likely to assume a fixed conformation distinct from that of the native molecule.

Anti-peptide antibodies can also be used for immunoaffinity purification of rare proteins⁷⁰. The main advantage of anti-peptide antibodies for immunoaffinity purification of proteins is that elution can be accomplished by excess peptide instead of the usual methods of high salt, low pH or chaotropic agents. Since the peptide elution is specific for the antigen-antibody union of interest, the recovered protein is less likely to be contaminated with extraneous proteins which were bound to the column at sites other than the antibody combining site. Also, proteins eluted by peptides are more likely to be recovered in a native state.

Theoretical considerations

The reason anti-peptide antibodies so often react with the native structure seems to be that they can adopt a number of conformations in solution, one of which approximates that adopted in the native molecule. If this is the case anti-peptide antibodies should contain a mixture of reactivities to these various conformations with only a small percentage of them actually reacting with the native structure. There are, however, other theoretical possibilities. One is that a protein molecule perturbed by solvent interactions exists as a statistical ensemble of conformational states with similar backbone structures but with variations in side-chain orientations on the surface of the molecule. In this

situation the anti-peptide antibody could either 'lock in' a particular conformation, induce a conformation, or, in fact, mould its own combining site which for similar reasons would itself be somewhat sterically mutable. That at least some of the latter notions may be true is supported by our studies on monoclonal antibodies to chemically synthesized peptides (H. Niman and R.A.L., unpublished). We selected monoclonal anti-peptide antibodies against an influenza virus HA1 peptide (amino acids 75-111) and then looked at what percentage would react with the native protein. Surprisingly, 14 of 27 reacted in an enzyme-linked immunoabsorbent assay (ELISA) and immunoprecipitation assays with intact HA1, a number much too large to be compatible with a simple model of occasional shared conformations between the free peptide and its counterpart in the protein. If these results can be extended and generalized, one will have to think in terms of flexible structures in protein, antibody molecules, or both. Alternatively, both in solution and as part of protein molecules, peptides may spend the bulk of their time in a limited number of comparable conformations presumably representing the state of lowest free energy.

Future prospects

The ability to make antibodies to defined regions of proteins will, no doubt, open the way to a number of biochemical experiments on the structure-function relationships of proteins as well as help in the immunological prevention and perhaps treatment of disease. Consideration of the theoretical basis for the immunogenicity of small synthetic peptides has led to questions about the fluidity of regions of protein molecules in solution, as well as the possibility that antibodies induce shape changes in proteins. The somewhat surprising results concerning the immunogenicity of peptides may lead to a better understanding of the dynamics of protein molecules in solution. The fact that such a high percentage of monoclonal anti-peptide antibodies react with the native protein already suggests that something unexpected is on the horizon.

I thank the many investigators who allowed me to cite their experiments before publication. I also thank my colleagues Greg Sutcliffe, Tom Shinnick, Nicola Green, Hannah Alexander, Steve Alexander, Richard Houghton, Jim Bittle, John Gerin, Bob Purcell, Dave Rowlands, Fred Brown, Henry Niman and Art Olson for their collaboration.

1. Sigal, N. H. & Klinman, N. R. *Adv. Immun.* **26**, 255 (1978).
2. Köhler, G. & Milstein, C. *Nature* **256**, 495 (1975).
3. Crumpton, M. J. in *The Antigens* (ed. Sela, M.) (Academic, New York, 1974).
4. Benjamini, E., Young, J. D., Shimizu, M. & Leung, C. Y. *Biochemistry* **3**, 1115 (1964).
5. LaPresle, C. & Durieux, J. *Ann. Inst. Pasteur* **92**, 62 (1957).
6. Cebra, J. J. *Immun.* **86**, 205 (1961).
7. Anderer, F. A. Z. *Naturforsch.* **18b**, 1010 (1963).
8. Atassi, M. Z. *Nature* **202**, 496 (1964).
9. Crumpton, M. J. & Wilkinson, J. M. *Biochem. J.* **94**, 545 (1965).
10. Landsteiner, K. *J. exp. Med.* **75**, 269 (1942).
11. Brown, R. K., Delaney, R., Levine, L. & van Vunakis, H. *J. biol. Chem.* **234**, 2043 (1959).
12. Goetzl, E. J. & Peters, J. H. *J. Immun.* **108**, 785 (1972).
13. Arnon, R. & Sela, M. *Proc. natn. Acad. Sci. U.S.A.* **62**, 163 (1969).
14. Arnon, R., Maron, E., Sela, M. & Anfinsen, C. B. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1450 (1971).
15. Parish, C. R., Wistar, R. & Ada, G. L. *Biochem. J.* **113**, 501 (1969).
16. Ichiki, A. T. & Parish, C. R. *Immunochimistry* **9**, 153 (1972).
17. Atassi, M. Z. & Saplin, B. J. *Biochemistry* **7**, 688 (1968).
18. Shinka, S. *et al. Biken's J.* **10**, 89 (1967).
19. Fujio, H., Imanishi, M., Nishioka, K. & Amano, T. *Biken's J.* **11**, 207 (1968).
20. Omenn, G. S., Ontjes, D. A. & Anfinsen, C. B. *Biochemistry* **9**, 313 (1970).
21. Sachs, D. H., Schechter, A. N., Eastlake, A. & Anfinsen, C. B. *J. Immun.* **109**, 1300 (1972).
22. Sutcliffe, J. G., Shinnick, T. M., Green, N., Liu, F.-T., Niman, H. L. & Lerner, R. A. *Nature* **287**, 801 (1980).
23. Min Jou, W. *et al. Cell* **19**, 683 (1980).
24. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366 (1981).
25. Green, N. *et al. Cell* **28**, 477 (1982).
26. Laver, W. G., Air, G. M., Dopheide, T. A. & Ward, C. W. *Nature* **283**, 454 (1980).
27. Webster, R. G. & Laver, W. G. *Virology* **104**, 139 (1980).
28. Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289**, 373 (1981).
29. Walter, G., Hutchinson, M. A., Hunter, T. & Eckhart, W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4882 (1981).
30. Walter, G., Scheidtmann, K.-H., Carbonek, A., Laudano, A. P. & Doolittle, R. F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5197 (1980).
31. Luka, J. & Klein, G. (in preparation).
32. Papkoff, J., Lai, M. H.-T., Hunter, T. & Verma, I. M. *Cell* **27**, 109 (1981).
33. Papkoff, J., Verma, I. M. & Hunter, T. *Cell* **29**, 417 (1981).
34. Sen, A. & Sen, S. (in preparation).
35. Wong, T.-W. & Goldberg, A. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7412 (1981).
36. Sen, A. & Sen, S. (in preparation).
37. Lukacs, G. & Baluda, M. (in preparation).
38. Robbins, K. C., Devare, S. G., Premkumar, R. & Aaronson, S. A. *Science* (in the press).
39. Shibasaki, T., Ling, N. & Guillemin, R. *Nature* **285**, 416 (1980).
40. Amara, S. G. *et al. J. biol. Chem.* **257**, 2129 (1981).
41. Schaffhausen, B., Benjamin, T. L., Pike, L., Casnellie, J. & Krebs, E. (in preparation).
42. Shinnick, T., Lerner, R. A. & Sutcliffe, J. G. (in preparation).
43. Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. *Nature* **293**, 543 (1981).
44. Green, N. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 6023 (1981).
45. Baron, M. H. & Baltimore, D. *Cell* **28**, 395 (1982).
46. Semier, B. L., Anderson, C. W., Hanecak, R., Dorner, L. F. & Wimmer, E. *Cell* **28**, 405 (1982).
47. Shinnick, T. M. & Blattner, F. (in preparation).
48. Berk, A. J., Lee, F., Harrison, T., Williams, J. & Sharp, P. A. *Cell* **17**, 935 (1979).
49. Jones, N. & Shenk, T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3665 (1979).
50. Graham, F. L., Harrison, T. & Williams, J. *Virol.* **86**, 10 (1978).
51. Jones, N. & Shenk, T. *Cell* **17**, 683 (1979).
52. Montell, C., Fisher, E. F., Caruthers, M. H. & Berk, A. J. *Nature* **295**, 380 (1982).
53. Feldman, L. & Nevins, J. (in preparation).
54. Goebel, W. F. *Nature* **143**, 77 (1939).
55. Goebel, W. F. *J. exp. Med.* **68**, 469 (1938).
56. Anderer, F. A. *Biochim. biophys. Acta* **71**, 246 (1963).
57. Langbeheim, H., Arnon, R. & Sela, M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4636 (1976).
58. Müller, G., Shapiro, M. & Arnon, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3403 (1982).
59. Lerner, R. A., Green, N., Alexander, H., Liu, F.-T., Sutcliffe, J. G. & Shinnick, T. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3403 (1981).
60. Dreesmann, G. R. *et al. Nature* **295**, 158 (1982).
61. Vyas, G. N. in *Hepatitis B Vaccine* (eds Naupas, P. & Guesry, P.) 227 (Elsevier, Amsterdam, 1981).
62. Prince, A. M., Ikram, J. & Hupp, J. P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 579 (1982).
63. Bittle, J. L. *et al. Nature* **298**, 30 (1982).
64. Sutcliffe, J. G. & Koprowski, H. (in preparation).
65. Gerin, J., Purcell, R. & Lerner, R. A. (in preparation).
66. Alexander, S., Alexander, H., Green, N. & Lerner, R. A. (in preparation).
67. Chou, P. Y. & Fasman, G. D. *Adv. Enzym.* **47**, 45 (1978).
68. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105 (1982).
69. Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824 (1981).
70. Walter, G., Hutchinson, M. A., Hunter, T. & Eckhart, W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4025 (1982).
71. Atassi, Z. & Lee, C. -L. *Biochem. J.* **171**, 429 (1978).
72. Blake, C. C. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. *Proc. R. Soc. B167*, 365 (1967).
73. Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. *Enzymes* **7**, 665 (1972).
74. Connolly, M. *Science* (submitted).

ARTICLES

High dynamic range mapping of strong radio sources, with application to 3C84

J. E. Noordam & A. G. de Bruyn

Radiosterrenwacht, Postbus 2, 7990 AA Dwingeloo, The Netherlands

A new reduction procedure is described to correct data obtained by the Westerbork Synthesis Radio Telescope. The technique allows very high dynamic range mapping of strong radio sources by also using redundant interferometers. The performance is illustrated with 6-cm and 21-cm continuum maps of 3C84, which have a dynamic range of about 10,000:1.

THE continuing improvement of radioastronomical detection systems during the past decade has resulted in the potential ability to detect ever fainter sources of radio emission. Aperture synthesis arrays in particular have been able to take advantage of the new low-noise receivers, to detect and map faint sources of radio emission. However, the synthesis arrays show increasing limitations when studying faint emission in the immediate vicinity of intense radio sources. These limitations are due to the corruption of the visibility amplitudes and phases of the correlated interferometer signals by ionosphere, troposphere, receivers and detectors. The way in which these errors, after Fourier inversion, will distort the source intensity distribution depends, among other things, on the array configuration. In the case of the Westerbork Synthesis Radio Telescope (WSRT) the map distortions tend to appear as (sections of) elliptical rings around and radial spokes emanating from the highest intensity features in the map¹. The dynamic range attained in a standard WSRT observation has previously been limited to 100–1,000, depending both on observing frequency and area of interest in the map. In special cases a technique called kneading² has been used to improve WSRT maps significantly over the result obtainable using standard reduction procedures.

Recently, several new data correction schemes^{3–5} have been introduced in aperture synthesis, and have produced high quality maps. In essence these schemes are all based on the assumption that errors can be traced to individual telescopes, rather than interferometers. In that case, the $N(N-1)/2$ interferometers one can form with N telescopes suffer from only N independent phase and gain errors per integration time (scan). The idea now is to try to adapt the data to a tentative model of the source intensity distribution, by allowing only the independent errors to vary, within specified limits. The fewer independent errors there are, the less freedom the data will have to comply with a wrong model, and it is assumed that the discrepancies will show the way to a better model for the next iteration step. This process has been known to converge to reliable and consistent results in many practical cases, although it seems to discriminate against extended structure. For the best results one needs many telescopes, and all telescopes

should participate in short as well as long baselines. The possibility of using redundant baseline information in the case of the WSRT constitutes an improvement in the sense that the number of independent errors can be reduced dramatically. This has opened the way to producing very high dynamic range maps ($>10,000:1$) of strong radio sources.

The method was tested—and to some extent developed—on the radio source 3C84, one of the strongest extragalactic radio sources at centimetre wavelengths, with total flux densities in 1981 (it varies) of 19 Jy at 21 cm and 60 Jy at 6 cm ($1 \text{ Jy} = 10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}$). Its emission is dominated by an unresolved inverted spectrum core⁶ whose intensity had previously prevented mapping at centimetre wavelengths the surrounding extended emission.

Use of redundant interferometers

The WSRT consists of 14 telescopes, 4 movable and 10 fixed, configured in a regular east–west array (Fig. 1). In the standard observing mode, 40 interferometers are formed by correlating signals from pairs of movable and fixed dishes. The advent of a new digital backend⁷ made it possible, in principle, to correlate all telescopes with one another, yielding 91 interferometers. However, in the present correlator, polarization information and a factor of 2 in sensitivity have to be traded to achieve this.

Although a minimum-redundant array produces more visibility information per telescope, the extra information provided by the redundant interferometers can be functionally exploited. When the array is positioned in its standard configuration (with separations $9A = AB = CD = 72 \text{ m}$, see Fig. 1), all spacings that carry redundant information, and a reasonably defined signal, can be compared to yield a weighted least-squares solution for all telescope errors. This reduces the number of independent errors per hour-angle scan (usually 2 min long) to only one gain and one phase error! (Because visibilities are compared, absolute intensity and position information are lost; the gain and phase stability of the WSRT, however, is such that these can usually be derived to 2% and 1 arc s, respectively, from the raw uncorrected data.) The corrected scans, which consist of complex visibilities on 38 different spacings, can be

Table 1 Summary of features of 3C84 structure and flux density

Component	Size		Flux density (Jy)	
	Angular	Linear*	1412 MHz	4995 MHz
VLBI core	$\sim 0.008 \times 0.016 \text{ arc s}$	$4 \times 8 \text{ pc}$	$13.2 \pm 0.2^\dagger$	59.6 ± 1.2
'30 arc s component'	$15 \times 40 \text{ arc s}$	$8 \times 21 \text{ kpc}$	2.3 ± 0.2	0.75 ± 0.05
Halo	$> 4 \times 6 \text{ arc min}$	$> 120 \times 180 \text{ kpc}$	> 2.4	?

* Assuming a distance of 106 Mpc, which is based on $V_{\text{sys}} = 5,300 \text{ km s}^{-1}$, $H = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

† In August 1980. One year later the core flux had increased to 14.4 Jy.

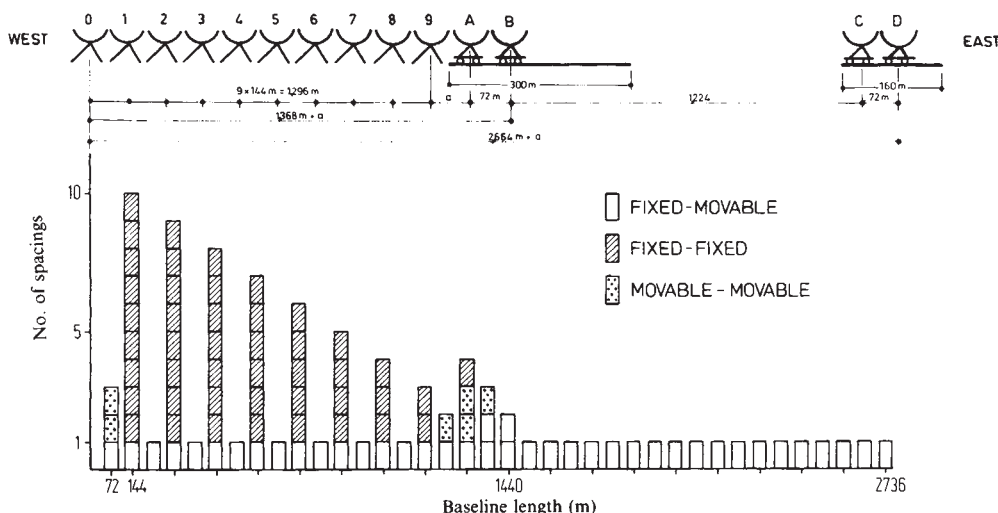


Fig. 1 Top, the geometry of the Westerbork array. The distance a between telescopes 9 and A is variable. In the optimum redundant configuration a is set to 72 m. Bottom, a histogram of the observed interferometer spacings when $a = 9A = 72$ m.

considered perfect internally and will be kept 'rigid' through-out further processing.

The decomposition into telescope errors also provides an r.m.s. measure of the consistency of the input data with the final solution. If telescope errors are indeed the only type of error, the level of inconsistency should be determined by the thermal noise. Down to a level of about 0.1° in phase and 0.2% in amplitude we find this indeed to be the case. Only when observing very intense sources, for example 3C84 at 6 cm, could amplitude and phase errors be detected at a level of 0.1° and 0.2% , which must be ascribed to individual interferometers. These errors, whose origin is still unknown, appear to remain constant during a 12-h observation and can therefore be removed in the reduction process (see below). These errors should be compared with the so-called closure errors in VLBI³ and MERLIN⁵ hybrid mapping.

Alignment of individual scans

Following the redundancy step the scans are perfect and 'rigid', by which we mean that from then on no adjustments of individual amplitudes or phases are allowed (or necessary). The need for alignment stems from several causes. The scans must be phase-aligned to correct for phase gradients over the array that are caused by short- and long-term atmospheric effects (troposphere at 6/21 cm, ionosphere at 21/49 cm). Amplitude-alignment is required to correct for slow variations in the overall system gain, extinction variations (6 cm) and scintillation (49 cm). Typical amplitude adjustments are in the range 1–3%; phase gradient adjustments can be as large as 50–100° across the array, but are more typically 10–20°.

We have investigated several ways of aligning scans. The preferred method would be to align without knowledge of the

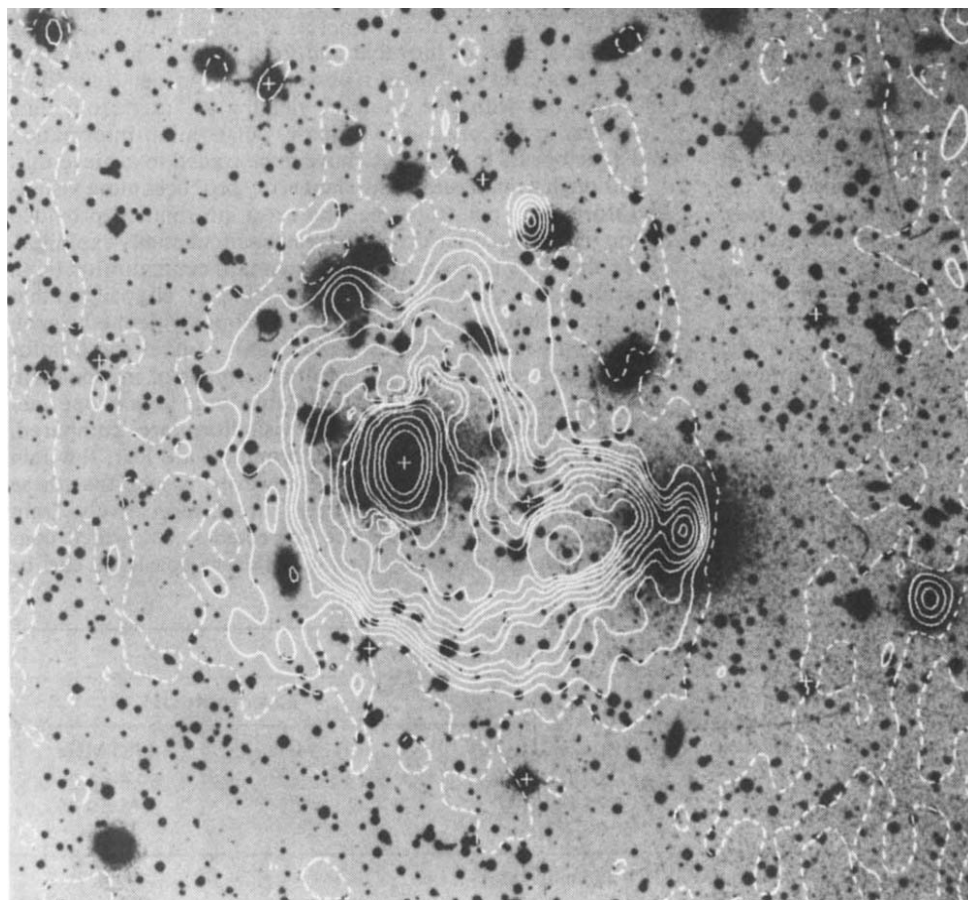


Fig. 2 The convolved 21-cm map of August 1980 superimposed on a deep optical photograph¹⁶. North is up and east to the left; the area shown measures 16.5×18 arc min. The resolution (30×45 arc sec FWHM) can be judged from the appearance of the point source in the upper-right quadrant. A point source of flux density 13.2 Jy has been subtracted at the position of the cross centred on NGC1275. The displayed contour levels are $-5, 0, 2.5, 10, 50, 100, 200, 400, 800, 1600, 3200$ and 6400 mJy per beam; the peak brightness is at 2,060 mJy per beam. The true zero level varies slightly across the map but lies at about -2.5 mJy in the inner 10 arc min of the map.

source structure. An elegant way of doing this is to use the fact that the centroid of an extended source should remain at a fixed location irrespective of the hour angle of the scan⁸. By shifting every scan to the projected position of this centroid one does not require a model of the source. A determination of the source centroid and total flux density using, for example, a one-dimensional CLEAN⁹, would suffice. Although this method is clearly easier for the user, the results obtained thus far are still inferior to those obtained using a model for the source. Nevertheless, the results are encouraging (dynamic range $\sim 1,000:1$) and for very extended complex sources, in particular, we expect it to become the preferred method.

At present, in a more conventional approach, the unaligned scans are Fourier inverted to produce a map of the synthesized field and then CLEAN⁹ is used to provide a simple model for the source. The individual scans are then adjusted to agree with the predicted model amplitudes and phases. In the alignment stage the user can specify the range of baselines over which he wants to force agreement between model and data. Depending on the source structure, for which some information can be gained by inspecting the visibility amplitudes, one could decide to align on the shorter or longer baselines. Following the alignment of the scans a better model for the source can be constructed and the alignments refined, if necessary. Here the method is similar to the self-calibration procedures used for VLA and MERLIN data. However, in the case of the WSRT only two parameters per scan need be determined. This, and the near-perfection of the individual scan complex visibilities, are the reasons that the process usually converges in one or two iterations. Finally, corrections can be applied for interferometer errors, if they are important (on sources stronger than several Jy). Following the final alignment corrections the difference between the model visibilities and the observed visibilities, averaged over 12 h, are determined for each interferometer. An incomplete model would only produce slowly varying differences as a function of interferometer baseline length. The misbehaving interferometers can then easily be spotted and corrected.

We have applied the technique to more than 12 sources at wavelengths of 6, 21 and 49 cm. The very high dynamic range attained in nearly all of these cases can, in summary, be attributed to mainly two factors: (1) the excellent design and performance of the correlator, limiting the magnitude of inter-



Fig. 4 The full resolution 6-cm map made from the April 1981 data is shown superimposed on the narrow band H_{α} photograph taken by Lynds²¹ (courtesy Dr Lynds and the Kitt Peak National Observatory). An unresolved source of 60 Jy was removed at the peak position. The residual peak is about 110 mJy. Contours are $\pm 6, 12, 18, 24$ (12) and 108 mJy per beam (3.5×5.3 arc s FWHM). North is up and east to the left; the region shown measures about 190×150 arc s.

ferometer-based errors; (2) the use of redundant baselines to reduce to only two the number of independent 'errors' per scan.

21-cm results on 3C84

3C84 is associated with NGC1275, the brightest galaxy in the Perseus cluster. Its radio structure is unusually complex and has been mapped previously with the Cambridge and WSRT arrays¹⁰⁻¹³. Two 12-h syntheses were made of the source in the redundant configuration at both 6 and 21 cm wavelength. The 21-cm observations were done on 14 August 1980 and 21 August 1981, and the 6-cm observations on 21 April 1981 and 2 January 1982. In the first 21-cm and both 6-cm observations the system temperature was ~ 90 K and the bandwidth 10 MHz. Feed polarization was linear in position angle 0° . During the more recent 21-cm observation the four movable telescopes were equipped with cooled low-noise receivers¹⁴, resulting in an overall system temperature of 55 K. The amplitude scale was calibrated using 3C286 as the primary flux density standard; it is assumed to have flux densities of 7.41 Jy at 4,995 MHz and 14.77 Jy at 1,412 MHz, which are on the Baars *et al.*¹⁵ scale. The coordinates in the various maps were tied to that of the core source in 3C84: RA = 03 h 16 min 29.57 s, dec. = $+41^{\circ}19$ arc min 51.9 arc s (1950.0).

Figure 2 shows the 21-cm total intensity distribution superimposed on a deep optical photograph¹⁶ of the central part of the Perseus cluster. A smoothed map is shown to bring up the extended halo around NGC1275. The map shows at much greater sensitivity and higher resolution the various structures seen previously^{10,13}. It also revealed the presence of weak discrete radio sources associated with NGC1270, NGC1272 and NGC1278. The dynamic range in this map varies from $\sim 3,000$ close to the source peak to 10,000 further out. 3C84 has three distinct radiation components of widely different angular scales (see Table 1). The most intense one is an unresolved VLBI core which was subtracted from the map. Underneath the core lies the so-called 30 arc s component which we find to be symmetrically extended, contrary to the findings of Miley and Perola¹⁰.

The diffuse emission surrounding the 30 arc s component has a surface brightness more than two orders of magnitude below that of the component itself. The relationship between these two components is unknown. This third component, which we will refer to as the halo, has no apparent boundary. At the lowest significant contour in our map the size is nearly 10 arc min. The brighter part of it, 4×6 arc min is elliptically shaped

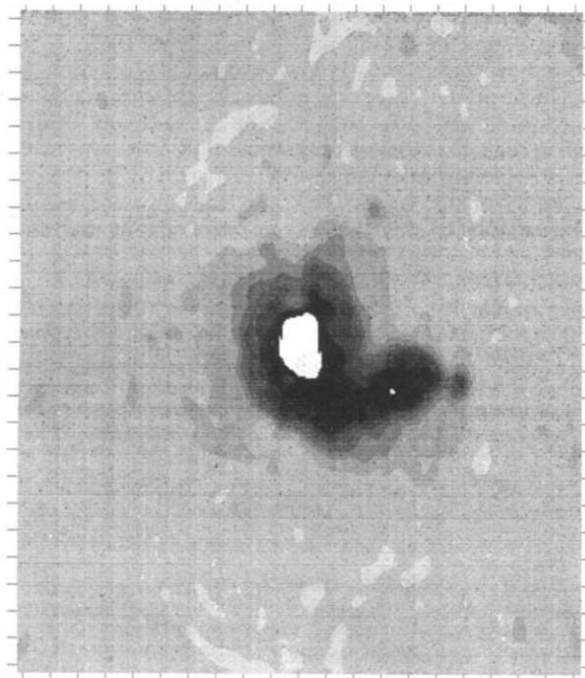


Fig. 3 Grey-scale presentation of the convolved 21-cm map shown in Fig. 2.

and centred on NGC1275. Since our shortest interferometer spacing (72 m) has a fringe separation of 10 arc min we cannot detect any more extended structure. Structure on a scale somewhat greater than 10 arc min was seen at 408 MHz (ref. 12).

An interesting feature of the halo is the bright region about 3 arc min south-west of NGC1275 of which the outermost contours enclose the galaxy NGC1272 which is a radio source in its own right. NGC1272 seems to be connected to NGC1275, through this feature, in a curved bridge of emission. However, we regard this connection as illusionary and attribute it to the superposition of a trail of emission associated with NGC1272 and a symmetrical halo surrounding NGC1275. Gisler and Miley¹³ have suggested that NGC1272 and NGC1275 are a pair of head-tail sources in orbit around each other. The fine structure within the region between the two galaxies, and the structure of the discrete source associated with NGC1272, do not support such a view. Below we discuss an alternative explanation for the enhanced emission features.

Various arguments suggest that NGC1272 is physically close to NGC1275, that is, not only in projection. If so, NGC1272 is moving through the composite thermal/non-thermal halo surrounding NGC1275. Such a trajectory is likely to affect the non-thermal radio emissivity (and, possibly also the thermal X-ray emissivity¹⁷). This scenario is strongly suggested by Fig. 3, which is a grey-scale representation of Fig. 2. A rough calculation supports this hypothesis quantitatively. The power deposited in the wake of a galaxy moving trans-sonically through the cluster medium is $\pi \rho v_g^2 R^2$, where ρ is the cluster density, V_g the galaxy velocity ($\sim 1,000 \text{ km s}^{-1}$) and R some characteristic radius^{18,19}. Where the radio brightness enhancement is highest, the density¹⁷ is $\sim 10^{-2} \text{ cm}^{-3}$. We take R to be 10 kpc, which is the accretion radius of NGC1272 for an adopted mass of $10^{12} M_\odot$. Although this cross-section is uncertain we note that the radio source associated with NGC1272 itself is also ~ 20 kpc diameter transverse to its hypothesized direction of motion. The total pathlength through the halo we estimate to be ~ 100 kpc. During the 10^8 yr taken by NGC1272 to traverse this distance, it deposits $\sim 10^{59}$ erg in the form of turbulent motions. How much of this energy eventually ends up in the form of magnetic and relativistic electron energy is unknown. The minimum (equipartition) energy contained in the 'trail' is $\sim 10^{57}$ erg and a 1% conversion efficiency therefore suffices to explain the emission from the trail. This does not seem implausible.

In the context of the wake hypothesis, the trail's emission might also be explained as a mere enhancement of the diffuse halo emission, without invoking *in situ* turbulent acceleration. As can be seen in Figs 2 and 3, the halo surface brightness at the location of the trail was non-zero before the passage of NGC1272. A moderate compression of the halo magnetic field in the wake of NGC1272 will lead to betatron acceleration of the relativistic electrons and with the enhanced field strength this will combine to yield a considerable increase of the synchrotron emissivity. However, the line-of-sight depth through the trail is considerably smaller than that through the halo and it is unlikely that this mechanism could explain the factor 30 enhancement in the surface brightness ~ 50 kpc behind NGC1272 where the wake reaches its peak brightness. If the wake hypothesis is correct, the observations provide direct evidence for turbulence-induced *in situ* particle acceleration in diffuse radio sources, on time scales of $\sim 5 \times 10^7$ yr.

The wake of NGC1272 shows some curvature; this is to be expected due to the dynamical pressure exerted by the gas in the accretion flow—which in the picture by Cowie *et al.*²⁰ rotates around NGC1275 with a period of $\sim 10^9$ yr—during the radiative (and kinematic) lifetime of the trail of $\sim 10^8$ yr. If NGC1272 can leave such an impressive mark, it is conceivable that part or all of the halo emission is due to continuous stirring through passages of other galaxies (see ref. 19). Another feature in the halo that might be interpreted as the remnant of a galactic wake is the spur in PA = 15° , extending out to 4 arc min. However, this is also the direction of the peculiar outlying H α filament²¹.

The 30 arc s component

Figure 4 gives the 6-cm map, which has nearly an order of magnitude higher resolution than the 21-cm map shown in Fig. 2. Although no trace of the halo could be found above a limiting flux density of about 2 mJy per beam area, this is not surprising in view of the steep spectrum¹³ of the halo and its resolved appearance at 21 cm. After subtraction of the intense unresolved core source the 6-cm map becomes dominated by the 30 arc s component referred to earlier. Our resolution is insufficient to decide whether there exists some emission on an arc second scale surrounding the core. The details of the structure within about one beamwidth from the core position therefore remain uncertain. The dynamic range in the 6-cm map varies from nearly 10,000:1 close in to about 50,000:1 further out. This was nicely demonstrated by the detection (independently in both 6-cm data sets), at a level of 4 ± 1 mJy, of the nuclear source in NGC1278, previously seen at 21 cm.

The 30 arc s component is oriented roughly in PA -10° , as first deduced by Miley and Perola¹⁰, and has an axial ratio of about 3:1. As has been noted before, this is also the direction of elongation of the core as deduced from VLBI studies^{6,22}. At the higher intensity levels the structure has some S-type symmetry. The overall dimensions of this component are roughly 15×40 arc s, corresponding to linear dimensions of 8×21 kpc. It is surprising that there is no structural relationship between the radio radiation and the H α filamentation system discovered by Lynds²¹ (Fig. 4). Fabian and Nulsen²³ suggested that the compression flow gave rise to the radio halo around NGC1275. Our 6- and 21-cm maps do not support this suggestion, considering the lack of correspondence between the halo emission and the optical filaments where the compression presumably is largest. However, the accretion flow may have profound effects on the structure of the 30 arc s component. If the rotating accretion flow continues its inflow beyond the stagnation point, which Cowie *et al.*²⁰ place at a radius of ~ 50 kpc, the dynamic pressure of the inflowing matter would eventually be strong enough to halt the relativistic particle flow from the nucleus, which is probably oriented in PA -10° . Transfer of angular momentum from the flow to the relativistic plasma clouds might then be responsible for the S-type symmetry in the 30 arc s component.

We thank Drs J. D. O'Sullivan, J. P. Hamaker, A. Bos and H. W. van Someren Gréve for informative discussions and assistance during the long period from planning to implementation of this reduction method. The Westerbork Synthesis Radio Telescope is operated by the Netherlands Foundation for Radio Astronomy (SRZM) with the financial support of the Organization for the Advancement of Pure Research (ZWO).

Received 28 May; accepted 2 August 1982.

- Hamaker, J. P. in *Image Formation from Coherence Functions in Astronomy* (ed. van Schooneveld, C.) 27–45 (Reidel, Dordrecht, 1979).
- Hamaker, J. P. in *Image Formation from Coherence Functions in Astronomy* (ed. van Schooneveld, C.) 47–53 (Reidel, Dordrecht, 1979).
- Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **223**, 25–36 (1978).
- Schwab, F. R. *Proc. 1980 Int. Optical Computing Conf.*, SPIE, **230**, 18–24 (1980).
- Cornwell, T. J. & Wilkinson, P. N. *Mon. Not. R. astr. Soc.* **196**, 1067–1086 (1981).
- Pauliny-Toth, I. I. K. *et al. Nature* **259**, 17–20 (1976).
- Bos, A., Raimond, E. & van Someren Gréve, H. W. *Astr. Astrophys.* **98**, 251–259 (1981).
- O'Sullivan, J. D. *NRAO Internal Note* 240 (1977).
- Högbom, J. A. *Astr. Astrophys. Suppl.* **15**, 417–426 (1974).
- Miley, G. K. & Perola, G. C. *Astr. Astrophys.* **45**, 223–226 (1975).

- Ryle, M. & Windram, M. D. *Mon. Not. R. astr. Soc.* **138**, 1–21 (1968).
- Birkinshaw, M. *Mon. Not. R. astr. Soc.* **190**, 793–799 (1980).
- Gisler, G. R. & Miley, G. K. *Astr. Astrophys.* **76**, 109–119 (1979).
- Casse, J. L., Woestenburg, E. E. M. & Visser, J. J. *IEEE Trans. Microwave Theory Tech.* **30**, 201–209 (1982).
- Baars, J. W. M. *et al. Astr. Astrophys.* **61**, 99–106 (1977).
- Arp, H. C. & Bertola, F. *Astrophys. J.* **163**, 195–201 (1971).
- Fabian, A. C., Hu, E. M., Cowie, L. L. & Grindlay, J. *Astrophys. J.* **248**, 47–54 (1981).
- Jaffe, W. *Astrophys. J.* **241**, 925–927 (1980).
- Roland, J. *Astr. Astrophys.* **93**, 407–410 (1981).
- Cowie, L. L., Fabian, A. C. & Nulsen, P. E. J. *Mon. Not. R. astr. Soc.* **191**, 399–410 (1980).
- Lynds, R. *Astrophys. J. Lett.* **159**, L151–L154 (1970).
- Unwin, S. C. *et al. Astrophys. J.* **256**, 83–91 (1982).
- Fabian, A. C. & Nulsen, P. E. J. *Mon. Not. R. astr. Soc.* **180**, 479–484 (1977).

Molecular structure of r(GCG)d(TATACGC): a DNA–RNA hybrid helix joined to double helical DNA

Andrew H.-J. Wang*, Satoshi Fujii*, Jacques H. van Boom†, Gijs A. van der Marel†, Stan A. A. van Boeckel† & Alexander Rich*

* Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Gorlaeus Laboratory, University of Leiden, 2300 RA Leiden, The Netherlands

The molecule r(GCG)d(TATACGC) is self-complementary and forms two DNA–RNA hybrid segments surrounding a central region of double helical DNA; its molecular structure has been solved by X-ray analysis. All three parts of the molecule adopt a conformation which is close to that seen in the 11-fold RNA double helix. The conformation of the ribonucleotides is partly determined by water molecules bridging between the ribose O2' hydroxyl group and cytosine O2. The hybrid–DNA duplex junction contains no structural discontinuities. However, the central DNA TATA sequence has some structural irregularities.

THE sugar–phosphate backbones of DNA and RNA differ in one fundamental feature—the presence of a 2' hydroxyl group in the RNA backbone, which makes it more reactive and less flexible than the DNA backbone. Hybrid molecules are formed when DNA directs the polymerization of RNA or when RNA is the template for DNA polymerization. The former process occurs during transcription and the latter is found in the action of reverse transcriptase. In DNA replication, Okazaki fragments are initiated by a short primer containing a few ribonucleotides¹. In all these cases, a hybrid helix is found containing an RNA and a DNA strand. Double helical DNA and RNA usually form different types of helices, the B form being common for DNA, and the A form for RNA. Thus, we are interested in understanding how these two types of backbones are accommodated in a hybrid molecule. The formation of a DNA–RNA hybrid helix was first demonstrated with synthetic polynucleotides². Advances in synthetic technique have allowed the synthesis of oligomers which can provide structural information about molecules of this type. In considering this problem, we decided to synthesize a self-complementary decamer containing three ribonucleotide units at the 5' end linked to seven deoxynucleotide units. This can form a molecule which has a DNA–RNA hybrid in the three base pairs at either end of the molecule and a central segment of four base pairs which is a DNA–DNA duplex. The DNA duplex and its attached DNA–RNA hybrid may serve as a model for visualizing the structure of the Okazaki DNA fragment after initiation of replication, as only a small number of ribonucleotides are used as a primer. Here we report crystallization of the decamer r(GCG)d(TATACGC) in a form suitable for X-ray diffraction analysis. We have determined the structure of the crystal and find that the hybrid molecule forms a helix which is of the A type or close to 11-fold RNA³. Further, the presence of an A-type helix in the hybrid portion of the molecule seems to have influenced the conformation of the central four base pairs of DNA, which also assume the same A-DNA conformation.

Crystallization and structure determination

Synthesis of the hybrid decamer has been described previously⁴. The decamer was crystallized in a solution containing 1.5 mM decamer, 30 mM sodium cacodylate buffer (pH 6.0), 8 mM spermine, and 15 mM magnesium chloride; 20 µl of this solution were placed in the depression of a glass spot plate, which was then enclosed in a sealed box containing a reservoir with 40% 2-methyl-2,4-pentanediol. Crystals were formed by vapour diffusion. After 1 week, thick chunky rods (0.4 × 0.4 × 0.7 mm) were observed growing in the solution. The crystals

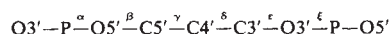
were mounted in sealed glass capillary tubes with a droplet of mother liquor and were found to diffract to a resolution of 1.9 Å. The space group was P2₁2₁2₁ with $a = 24.20$ Å, $b = 43.46$ Å and $c = 49.40$ Å. The asymmetric unit consisted of one decamer duplex containing 20 bases. Data were collected on a Nicolet diffractometer to 2 Å resolution at 0 °C.

Of 3,798 independent reflections, 2,521 were found to be observable over the 1.5 σ (I) level. A rotation–translation search was attempted using A-DNA as a model⁵. This search produced only moderate agreement using the data from 13 to 5 Å resolution, and the model would not refine further. We believed this might be due to the fact that the molecule has slightly different conformations in the three different segments of the molecule, that is, the two DNA–RNA hybrid segments and the DNA–DNA duplex. Another search was carried out using a model with three different segments covering the two hybrid parts and the DNA duplex. This refinement strategy did not converge, therefore another model was used in which the molecule was broken up into three segments containing four base pairs at either end and two in the middle. An 11-fold

Table 1 Averaged torsional angles* and helical parameters

	DNA–RNA hybrid (this work)	A-DNA ³ (fibre)	A-RNA ³ (11-fold) (fibre)	B-DNA ³ (fibre)
Torsional angle (deg)				
α	–69(31)	–90	–60	–39
β	175(14)	211	175	209
γ	55(22)	47	49	31
δ	82(9)	83	83	157
ϵ	–151(15)	–185	–147	–201
ξ	–75(16)	–45	–78	–99
χ	–162(10)	–153	–167	–95
P	17(15)	14	14	192
Helical parameters				
Rotation per unit (deg)	33(3)	33	33	36
Rise per unit (Å)	2.6(3)	2.6	2.8	3.4
Residues per turn	10.9	11.0	11.0	10.0
Propeller twist angles (deg) in base pairs				
Dihedral angle	14	11	12	1
Roll component	11	–11	–12	1
Tilt component	–2	0	0	0
Adjacent phosphorus atom separation (Å)				
	5.9(2)	5.7	5.7	6.5

Torsional angles are defined as



and χ is the glycosyl torsion angle. P is the phase angle of pseudorotation¹⁹.

* End effects are excluded and the standard deviations are given in parentheses.

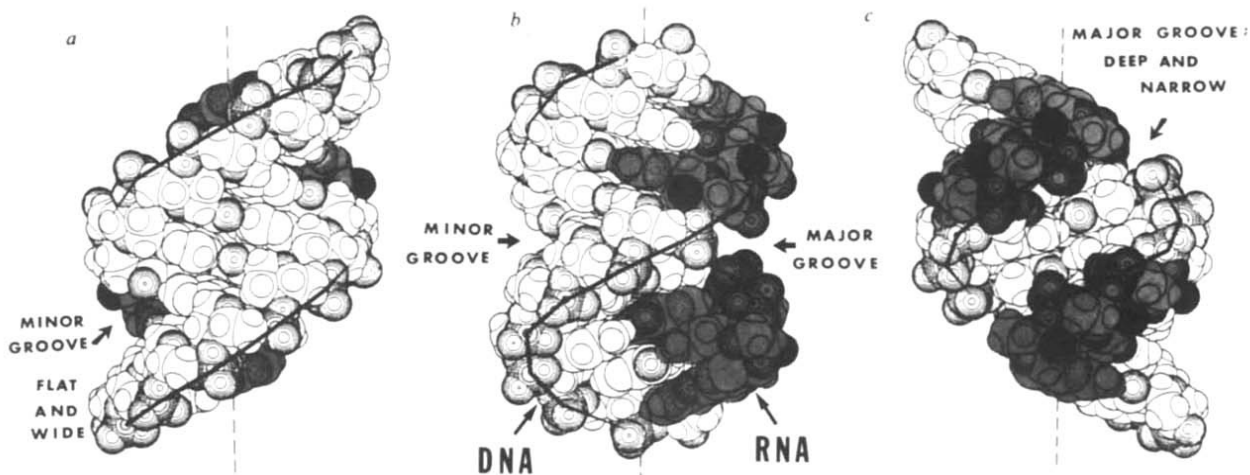


Fig. 1 A space-filling drawing of the molecule is shown in three different orientations 90° apart from each other about the dashed vertical axis. *a*, Looking into the minor groove of the molecule which is broad and flat. The phosphate groups on opposite strands are 8.6 Å apart across the minor groove. The heavy black lines are drawn from phosphate to phosphate to show the flow of the polynucleotide chain. The ribonucleotides are shaded and the ribose 2' oxygen atoms are solid. *b* Diagram shows both the minor groove on the upper part of the molecule and the major groove at the lower right. The shaded ribonucleotide backbone segments on the right are close to each other on either side of the major groove. The bases are tilted 19° from the vertical axis. In this centre view, the tilt reverses itself in going from the upper part to the lower part of the molecule. *c*, Looking into the deep narrow major groove of the molecule. The phosphate groups are 4.7 Å away from each other across this groove. Solid circles, oxygen atoms; stippled circles, nitrogen atoms; spiked circles, phosphorus atoms.

RNA model was then used for all three segments and an additional search was carried out to find local minima. This search was satisfactory, producing an *R* factor of 36% at 5 Å resolution. The model was further refined using the Konnerth-Hendrickson restrained refinement program⁵ to a final *R* factor of 0.16. As the molecule gradually refined, a large number of water molecules appeared in the electron density map; at the end, 179 water molecules could be located surrounding each decamer. However, it was not possible to make an unambiguous assignment for either spermine molecules or Mg²⁺ ions in the electron density map even though they have been observed in other oligonucleotide structures which diffracted to higher resolution⁶.

Description of the molecule

Although the asymmetric unit is a 10-base pair (bp) duplex, the molecule possesses a non-crystallographic 2-fold axis with a r.m.s. deviation of 0.95 Å between the two strands of the duplex, by least-square fitting. The two strands thus have almost identical conformations. The molecule has an almost perfect A-type geometry with all the ribose and deoxyribose sugars in the C3' *endo* conformation; this is the conformation normally found in ribonucleotides^{7,8} and in the A-form of DNA³. Three different views of the molecule are shown as van der Waals' diagrams in Fig. 1. To aid in the identification of the ribonucleotides, they have been shaded and the ribose 2' oxygen atoms are black. Figure 1*b* shows both the minor groove on the left side of the molecule and the major groove on the right side. The views of either side of the molecule are obtained by a rotation of 90° about the vertical axis. At the right we are looking into the deep, narrow major groove that has a depth of 8 Å. The view at the left shows the broad, flat minor groove with a depth of only 2–3 Å. The difference between these two surfaces is quite great in the A conformation in contrast to the B-DNA conformation where the two grooves are more equivalent.

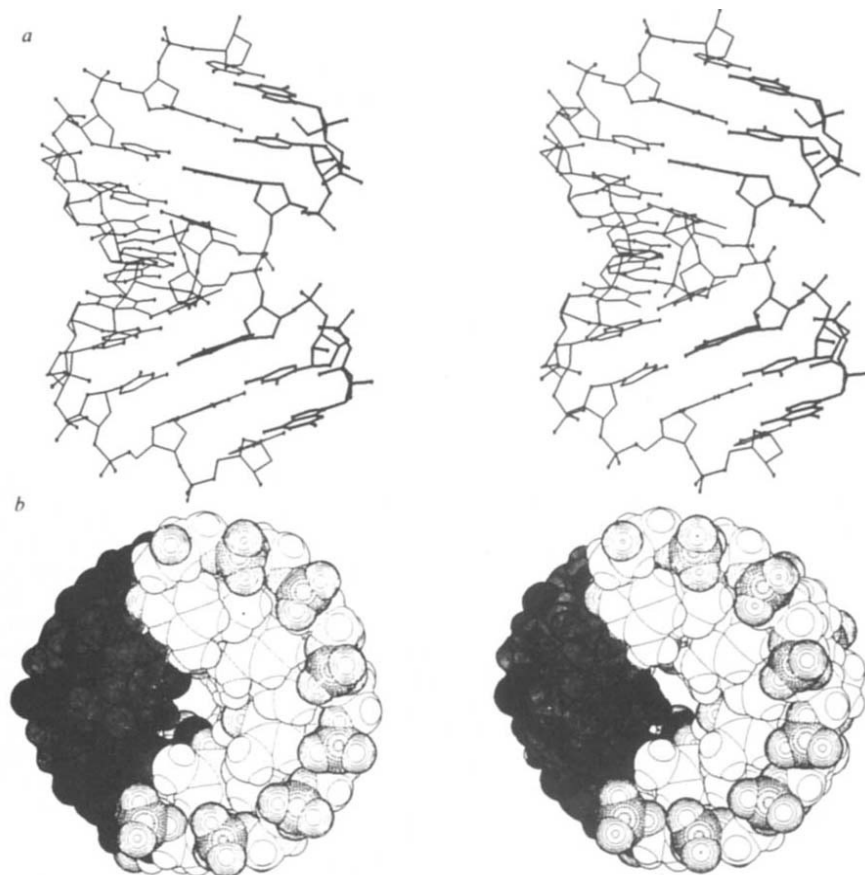
Figure 2*a* is a stereo skeletal side view of the molecule. The base pairs have an average rotation per residue of 33° and a rise along the axis of 2.6 Å. In this helix there are 10.9 residues per turn; we are thus observing just under one complete turn. The base pairs are considerably removed from the axis of the molecule and are tilted at an angle of 20° to the helix axis. Three different crystal structures of DNA oligonucleotides have been reported that assume A-DNA conformations: two octamers^{9,10} and one tetramer¹¹. However, the helical param-

eters of the present molecule are closer to that of the A-type geometry which has been studied extensively in both DNA and RNA fibres (Table 1). This may reflect the influence of the hybrid portions of the molecule. It is interesting that the perfection of the A-type geometry is found not only in the hybrid sections but also in the 4-bp central DNA section. The space-filling stereo view in Fig. 2*b* shows the molecule as viewed down the axis. There is an empty space nearly 2 Å in diameter along the axis of the molecule.

The Watson-Crick base pairs all have a propeller twisting (see Fig. 2*a*), the average dihedral angle of which is 14° excluding the two terminal base pairs which are involved in the intermolecular interactions. The sign of the propeller twisting between the base pairs is the same as that observed in a crystal of a B-DNA dodecamer¹². We have examined the hydrogen bond lengths in the various segments of the molecule. The four A-T base pairs in the DNA-DNA segment of the centre of the molecule have average bond lengths for NH—N of 2.8 Å and for NH—O, 3.12 Å. The six G-C base pairs in the DNA-RNA hybrid portion of the molecule have average hydrogen bond lengths: N—HN, 2.86 Å; N4H—O6, 2.84 Å; and N2H—O2, 2.80 Å. The nucleotide conformations are not identical but fall within a fairly narrow range. The average torsion angles for the sugar-phosphate backbone are listed in Table 1. A detailed listing of the torsion angles for the individual residues will be published elsewhere.

There is considerable interest in the contribution made by the 2' hydroxyl group of the ribose residues to the overall conformation. The 2' hydroxyl stabilizes the C3' *endo* conformation of the sugar¹³. Several model building proposals have been made which involve the ribose 2' hydroxyl group in different types of intramolecular hydrogen bonding to adjacent nucleotides in the chain¹⁴. We have examined the six hydroxyl groups present in this structure. The two hydroxyl groups associated with the pyrimidine residues have a hydrogen bonding interaction with a water molecule located 2.93 Å away. The water is hydrogen-bonded (3.07 Å) to O2 of the same cytosine residue (Fig. 3). The 2' hydroxyl group thus seems to play a significant part in stabilizing the conformation in the ribonucleotide segments of the molecule through the use of a bridging water molecule. Examination of the hydroxyl groups found in the four riboguanosine residues reveals that they also are hydrogen-bonded to water molecules, but the latter are hydrogen-bonded to adjacent molecules in the lattice. An intramolecular bridging water molecule similar to that seen on

Fig. 2 Stereo views of the hybrid helix. *a*, The molecule in the same orientation as that seen in Fig. 1*b*. The ribonucleotide segments are drawn with a thicker line than the deoxyribonucleotide segments. In this stereo view with a skeletal backbone, it can be seen that the base pairs have a propeller twist between them. The propeller twist angle has an average value of 14° throughout the structure, excluding the terminal base pairs. *b*, Stereo view of the molecule drawn with space-filling atoms viewed down the helix axis. The axis of the molecule has an empty space ~ 2 Å in diameter which is not occupied. The base pairs form a prominent right-handed staircase-like arrangement in this view; the phosphate groups have a conformation such that they are tipped into the major groove of the helix, so that they form an interrupted 'railing' for the staircase. The shading of the atoms and the nucleotides is the same as for Fig. 1.



the cytosine residues has been found between the O2' hydroxyl group and the N3 of guanine in the crystal structure of the dinucleoside monophosphate GpC⁸. It is perhaps reasonable to believe that such bridging water molecules could be a constant feature of RNA double helices in solution. Note that bridging water molecules have been found in several nucleic acid structures where they are believed to have a significant role. For example, a water molecule is found bridging between guanine N2 and the phosphate oxygen of the next nucleotide in the left-handed Z-DNA helix⁶; this water molecule stabilizes the *syn* conformation of the deoxyguanosine residue. Another example is found in the interaction of the antibiotic daunomycin with a DNA oligonucleotide, where the acetyl group of the daunomycin molecule is linked to a guanine residue in the base pair adjacent to the intercalative site through the use of a bridging water molecule¹⁵.

Studies have been made of the exchange rate of protons attached to the 2' hydroxyl group of ribonucleotides in a DNA-RNA hybrid model system¹⁶. These protons have a slow exchange rate, which suggests that they might be involved in some kind of stable structural interaction. Interactions with bridging water molecules may be responsible for some of this behaviour.

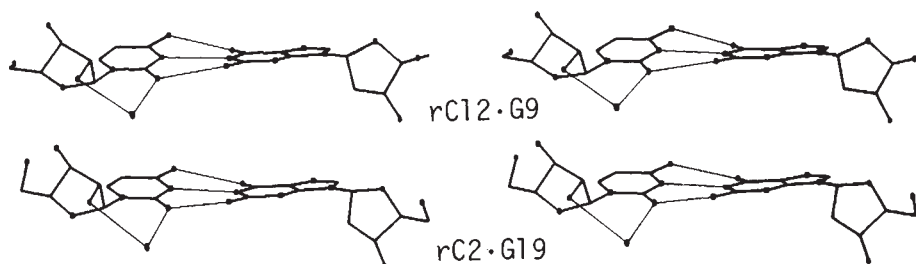
The stacking of successive nucleotides along the helix is similar to that found in a typical A-DNA helix or the similar 11-fold RNA helix³. The stacking interactions at the hybrid-deoxynucleotide interface were found to be similar to those observed in purine-pyrimidine sequences elsewhere in the

molecule. Thus no structural discontinuity is found when the stacking is examined in detail. Note that there are some bends in the helix axis at the TpA steps. The dihedral angles between the TATA base pairs show an interesting discontinuity. The angle between the first two base pairs is 19° , while that of the second step is 16° . In both cases, the opening of the base pairs is always in the direction of the minor groove. This explains the fact that in the structure determination, we obtained a satisfactory rotation-translation search only when the molecule was divided into 4-2-4-bp segments. Note also that the two central A-T base pairs have elongated NH—O hydrogen bonds with lengths averaging 3.23 Å. The distortions associated with the TATA sequence observed in this structure may be a sequence-dependent property of the DNA molecule, and is noteworthy in view of the special role which TATA sequences have in the initiation of transcription by RNA polymerase.

Lattice packing

The lattice is rather tightly packed and contains 33–35% water by mass, an average value for oligonucleotide crystals. The molecules are organized around a series of 2-fold screw axes running through the lattice in three perpendicular directions. The major interactions between the molecules are due to the flat base pairs at both ends which stack on adjacent molecules in the broad, shallow minor groove. The nucleotides are numbered from the 5' to the 3' end on both chains, so that rG1 is paired to C20. Stacking of the terminal base pairs is shown in

Fig. 3 Stereo skeletal drawings of two different guanine-cytosine pairs in the hybrid helix. (The nucleotides are numbered 1–10 on one chain, 5'→3', and 11–20 on the other chain in the same 5'→3' direction.) The figures show a similar conformation of the cytosine ribose residues at the left in both base pairs. A bridging water molecule at the lower left in both diagrams is hydrogen-bonded both to the 2' hydroxyl group of the ribose and the O2 of cytosine. The electron density of the water molecule indicates that it is fully occupied, thus it is possible that this bonding helps to stabilize the orientation of the base relative to the sugar.



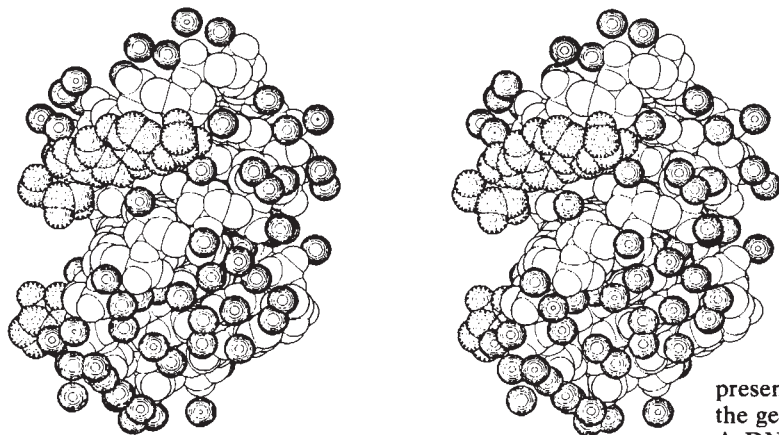


Fig. 4. The base pair rG1·C20 abuts onto nucleotides A7 and C8 of an adjoining molecule. The guanine base of rG1 makes van der Waals' contacts with the adenine A7 (atoms N3, C1' and C2') while the cytosine ring of C20 makes close contacts with the C1' and C2' atoms of the sugar ring of cytosine C8. The long direction of the base pair rG1·C20 lies along the path of the minor groove (Fig. 4).

The contacts at the other end of the molecule involving C10·rG11 have a similar orientation but they differ in detail. A hydrogen bond is found between the oxygen O3' of C10 and N2 of guanine G19. This hydrogen bond may be important in building up the lattice because of its directionality. Another interaction involves the O2' oxygen of rG11 which hydrogen bonds to a bridging water molecule, which in turn donates a proton to atom N3 of the adenine ring on A17. Other close van der Waals' contacts are found between the guanine ring of rG11 and the sugar phosphate chain of C18. The interaction at this end of the molecule is associated with hydrogen bonding as well as a bridging water molecule, in contrast to the strictly van der Waals' interactions found at the opposite end.

The major motif in building up the lattice is the stacking of planar base pairs on the flat minor grooves. The importance of this type of interaction in the lattice formation of A-DNA fragments has been discussed previously⁹. Similar interactions are found in three other lattices which form A-type helices, including the DNA octamers d(GGCCGGCC)⁹ and d(GGTATACC)¹⁰, and the tetramer d(CCGG)¹¹. The prevalence of these interactions in building up lattices is probably due to the fact that the energy is minimized by having the hydrophobic planar purine-pyrimidine base pair stacked on the relatively hydrophobic planar A-DNA minor groove. Neither the B-DNA helix, with a deeper minor groove, nor Z-DNA has a planar surface of this type, thus different types of lattice packing interactions are found in these cases. In Z-DNA crystals⁶ and in the daunomycin-B-DNA crystal complex¹⁵, the planar purine-pyrimidine base pairs at the ends of the molecules stack on themselves to build up a lattice.

Oligonucleotides as models

The mating of ribonucleotide strands with deoxyribonucleotide strands to form a hybrid double helix is one of the principal reactions of the biological information transfer system. DNA-RNA hybrid molecules are found covalently linked to a DNA duplex at the initiation of DNA replication in the Okazaki fragments¹ and are also found in mitochondrial DNA¹⁷. The

Fig. 4 A space-filling stereo diagram of the molecule showing the water molecules (solid circles) which are found in the first hydration sphere. In addition, base pairs from the ends of other molecules are shown stacked in the minor groove (interrupted circles). Fewer water molecules are found in the flat minor groove than in the major groove.

present structure has allowed us to visualize for the first time the geometry of the hybrid helix. We find that the helix has an A-DNA or 11-fold RNA geometry. It consistently uses C3' *endo* puckering of both ribo- and deoxyfuranose rings, so that adjacent phosphate groups are closer together than they are in the B-DNA double helix which uses C2' *endo* puckering. It seems that the RNA backbone has forced its geometry on the DNA backbone in a stable conformation. This may be due to the fact that deoxynucleotides can have either ring pucker, but ribonucleotides favour the C3' *endo* pucker.

It should be noted that the 2' oxygen atoms in the ribonucleotide segment are prominent on the outer surface of the molecule (Fig. 1). As the overall conformation of the residues of both the ribose and the deoxyribose residues is almost the same in the hybrid molecule, the presence of the hydroxyl group on the surface means that proteins or other molecules interacting with this can readily detect the difference between an RNA fragment forming an A-helix and a deoxynucleotide segment having the same conformation.

The central four base pairs of the DNA-DNA duplex also have the A-DNA geometry usually found in the RNA double helix; possibly this is due to the influence of the hybrid segments at either end of the molecule. If this is the case, it would be interesting to know how long a stretch of DNA would be influenced by an A conformation at one point. The answer to this would elucidate the nature of the A-to-B transitions which undoubtedly occur at many points in long DNA molecules.

Another question in this regard is whether the central 4-bp DNA duplex of this molecule would retain an A-DNA conformation in solution. A NMR study of this molecule in solution suggests that it does retain the A-DNA conformation, similar to that seen when the molecule is in the crystal lattice¹⁸. The latter study supports the idea that the A conformation must be due to the ribonucleotide components because the same molecule composed entirely of deoxyribonucleotides shows the typical B-DNA conformation.

The availability of pure fragments of oligonucleotides constructed using both the ribo and the deoxy series may help to further our understanding of the chemical forces which stabilize different conformations. As different conformations are often used in different ways in biological systems, these studies may eventually provide a firm basis for understanding the manner in which the nucleic acids perform their functions.

We thank Dr Gary Quigley for helpful discussions and for the use of his graphic plotting programs. This research was supported by grants from the NIH, NASA, the American Cancer Society and the Netherlands Organization for the Advancement of Pure Research (ZWO).

Received 7 July; accepted 12 August 1982.

- Ogawa, T. & Okazaki, T. A. *Rev. Biochem.* **49**, 421-457 (1980).
- Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **46**, 1044-1053 (1960).
- Arnott, S., Smith, P. J. C. & Chandrasekaran, R. *CRC Handbk Biochem. molec. Biol.* 3rd edn, Vol. 2 (ed. Fasman, G. D.) (CRC, Cleveland, 1976).
- van Boeckel, C. A. A. *et al. Recl Trav. chim. Pays-Bas Belg.* **100**, 389-390 (1981).
- Hendrickson, W. A. & Konner, J. *Biomolecular Structure, Conformation, Function and Evolution* Vol. 1 (ed. Srinivasan, R.) 43-57 (Pergamon, Oxford, 1979).
- Wang, A. H.-J. *et al. Nature* **282**, 680-686 (1979).
- Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. *J. molec. Biol.* **104**, 109-194 (1976).
- Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. *J. molec. Biol.* **104**, 145-167 (1976).
- Wang, A. H.-J., Fujii, S., van Boom, J. H. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3968-3972 (1982).
- Shakked, Z. *et al. Proc. R. Soc. B213*, 479-487 (1981).
- Conner, B. N., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. *Nature* **295**, 294-299 (1982).
- Wing, R. M. *et al. Nature* **287**, 755-758 (1980).
- Davies, D. B. *Prog. nucl. magn. Resonance Spectrosc.* **12**, 135-225 (1978).
- Ts'o, P. O. P., Rappaport, S. A. & Bollum, F. J. *Biochemistry* **5**, 4153-4156 (1966).
- Quigley, G. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7204-7208 (1980).
- Selsing, E., Wells, R. D., Early, T. A. & Kearns, D. R. *Nature* **275**, 249-250 (1978).
- Bernicke, A. & Clayton, D. A. *J. biol. Chem.* **256**, 10613-10617 (1981).
- Mellema, J.-R. *et al. Nature* (submitted).
- Altona, C. & Sundaralingam, M. *J. Am. chem. Soc.* **94**, 8205-8212 (1972).

LETTERS TO NATURE

Optical polarization position angle versus radio source axis in radio galaxies

Robert R. J. Antonucci

Lick Observatory, Board of Studies in Astronomy and Astrophysics, University of California, Santa Cruz, California 95064, USA

Optical polarization position angles tend to align with large-scale radio structure in low polarization quasars¹, indicating the existence of a simple geometrical relationship between the inner, optically-emitting region and the outermost, radio-emitting region. However, the meaning of the alignment is unclear because the cause of the optical polarization is unknown. If the polarization is caused by a synchrotron emission process then we are learning the direction of the magnetic field in the optically-emitting region. If it arises from scattering, then the position angle refers to the distribution of scatterers. At Lick Observatory we have been observing some quasars in the Stockman *et al.* sample¹ spectropolarimetrically to distinguish between the two possibilities (J. S. Miller and R. R. J. A. in preparation). If the polarization arises from the continuum emission process, only the continuum is polarized. If due to scattering, some of the line emission is probably polarized too. I have obtained optical spectropolarimetry data and VLA radio maps for a sample of radio galaxies to identify and interpret such alignment effects. The entire data set will be published as part of a study of the relationship between optical and radio properties of radio galaxies (work in preparation). The results show that most radio galaxies with optical nuclear emission are in fact polarized. There is a population showing quasar alignment effect, and unexpectedly there is also a population showing a perpendicular relationship. Emission-line polarizations exhibit various behaviours but in all cases make it very likely that the polarizations of both groups arise from scattering or dust transmission rather than the emission process itself.

Optical spectropolarimetry data for this project were obtained with the Lick 3-m Image Dissector Scanner as modified for simultaneous spectropolarimetry and spectrophotometry². Broadband polarizations presented here were obtained by integrating the spectropolarimetry data. Objects were chosen for polarization measurement by brightness and emission-line strength, consistent with the need to include a variety of optical and radio morphological types.

In most cases the radio position angles were measured from 6-cm radio maps made by E. B. Fomalont and myself with the NRAO Very Large Array. (The exceptions are 4C31.32 (ref. 3) and 3C234 (ref. 4).) The VLA instrument and standard reduction programs have been described previously⁵.

Table 1 shows the absolute value of the difference between the position angle of optical polarization and that of large-scale radio structure. It includes all those with constant polarization position angle and apparently not in danger of contamination by galactic interstellar polarization⁶. The result is that an aligned population analogous to the quasars is present, and unexpectedly there is apparently a group showing a perpendicular relationship. Note from Table 1 that the statistical errors in the polarization position angle are sometimes large; the best observed objects generally show the most precise alignments. For the low-polarization objects, there may be additional errors due to superposed galactic interstellar polarization. The uncertainties in measuring the radio axis position angles are estimated to be 5°.

As a test of the reality of the aligned group, the cumulative binomial probability distribution was computed. A random

distribution of position angle differences would result in at least five objects in the range 0°–20° only 5% of the time.

A similar analysis of the 90° peak does not yield a very high level of significance, but its reality is supported by (1) the fact that the best-observed objects in the peak show the most precise perpendicularity, and (2) the single significantly deviant object, 4C29.06, has an S-shaped radio map⁷, with the inner contours perpendicular to the optical polarization. More recent observations designed to determine whether the inner radio contours are more closely related to the optical polarization indicate that this is likely. Further observations to increase the sample size are under way.

Note that the most geometrically reasonable locations for peaks in the position-angle difference histogram are 0° and 90°. This is because any relationships other than alignment or perpendicularity would be somewhat dependent on viewing aspect and so would be washed out in an observational histogram. Also, in an axially-symmetric geometry, scattering polarization at intermediate angles would cancel out. Axially-symmetric geometries such as optically-thin and optically-thick scattering disks would produce parallel and perpendicular groups, respectively, independent of viewing aspect.

Referring again to Table 1, we see that the perpendicular group members tend to be N-systems. They are generally more highly polarized, more photometrically variable and systematically different in spectral properties when compared to the aligned group.

Spectropolarimetry data indicate that the polarization behaviour in the emission lines is varied for different objects and among the lines of a particular object. However, in all well-observed cases for both the parallel and the perpendicular subgroups, it appears that the polarization arises from a scattering or dust transmission effect rather than from the continuum-emission process. This is true even for the highly polarized object 3C234. In that object the broad H α line is polarized like the neighbouring continuum, while the forbidden [O III] lines have low polarization. The object is also unusual among active objects in that, although the polarization is very high (~10% in the continuum), its polarization is constant in time.

I will discuss the implications of these results elsewhere where the full radio, optical spectroscopic and optical spectropolarimetric data are available. To summarize the observations presented here: (1) most radiogalaxies with optical nuclei are polarized; (2) almost all well-observed objects have optical polarizations aligned with or perpendicular to the associated radio sources; (3) all polarizations are apparently due to scattering or dust transmission.

I thank Dr J. S. Miller for guidance throughout this project; this work was partially supported by NSF grant AST 80-19322. This is Lick Observatory bulletin no. 934.

Table 1 The polarizations are averages of two or more measurements

Object	Optical type	Optical polarization		Radio structure position angle	Difference
3C382	D3	1.29 ± 0.04%	58 ± 1°	54 ± 5	4 ± 5
4C29.41	ED	1.70 ± 0.36	85 ± 6	75	10 ± 8
3C287.1	N	3.10 ± 1.55	73 ± 14	88	15 ± 15
4C31.32	E	0.78 ± 0.30	143 ± 11	161	18 ± 12
3C305	Sa,pec	0.75 ± 0.18	94 ± 7	112	18 ± 9
4C02.39	N	2.11 ± 1.10	56 ± 15	117	61 ± 16
3C79	N	1.42 ± 0.78	172 ± 16	105	67 ± 17
4C29.06	E	1.92 ± 0.27	128 ± 4	57	71 ± 6
3C227	N	1.94 ± 0.28	29 ± 4	125	84 ± 6
3C234	N	7.20 ± 0.25	159 ± 1	68	89 ± 5

Some measurements vary in percent polarization. Objects which vary in polarization position angle have been excluded.

Received 24 May; accepted 12 August 1982.

1. Stockman, H. S., Angel, J. R. P. & Miley, G. K. *Astrophys. J. Lett.* **227**, 55–58 (1979).
2. Miller, J. S., Robinson, L. B. & Schmidt, G. D. *Publ. astr. Soc. Pacif.* **92**, 702–712 (1980).
3. van Breugel, W. J. M. *Astr. & Astrophys.* **81**, 275–281 (1980).
4. Riley, J. M. & Pooley, G. G. *Mem. R. astr. Soc.* **80**, 105–137 (1975).
5. Thompson, A. R., Clark, B. G., Wade, C. M. & Napier, P. J. *Astrophys. J. Suppl.* **44**, 151–167 (1980).
6. Mathewson, D. S. & Ford, V. L. *Mem. R. astr. Soc.* **74**, 139–182 (1970).
7. Katgent-Merkelijn, J., Lari, C. & Padrielli, L. *Astr. Astrophys. Suppl.* **40**, 91–118 (1980).

Distance to Crab-like supernova remnant 3C58

D. A. Green & S. F. Gull

Mullard Radio Astronomy Observatory, Cavendish Laboratory,
Madingley Road, Cambridge CB3 0HE, UK

The radio source 3C58 (G130.7+3.1) marks the probable position of the supernova of AD1181 (ref. 1). In its morphology, apparent size and radio spectrum, it is remarkably similar to the Crab Nebula (SN1054), being a 'filled-centre' supernova remnant (SNR), rather than the more usual shell-type. When comparing the physical properties of 3C58 with the Crab, or other 'filled-centre' remnants, a knowledge of its distance is required, and this has been a long-standing problem. Often, a value, or lower limit, of ~ 8 kpc has been assumed^{2,4}, but this implies that 3C58, although slightly younger than the Crab, is some five times larger. Also, it places 3C58 over 400 pc from the galactic plane, which is very surprising if it, like the Crab, is the remnant of a Type II supernova. The observations presented here show that 3C58 is in fact much closer, at a distance of only 2.6 kpc.

Previous estimates of the distance rest, like ours, on measurements of absorption by interstellar neutral hydrogen, but they all suffer from some drawback. The single-dish measurements² lack sufficient angular resolution to disentangle the complicated filamentary structures in the interstellar medium (ISM), whereas interferometric observations^{3,4} do not provide the information on the large-scale structure that is needed for reliable calibration of the spectrum. These studies have, however, shown that there is very little, if any, absorption ($\tau < 0.1$) of the 3C58 continuum for velocities beyond -35 km s⁻¹.

With this in mind, we observed 3C58 with the Cambridge Half-Mile Telescope (HMT) at 21 cm. Two complete surveys, overlapping in velocity coverage, were made in November 1981. Details of the telescope, receivers and surveys are given in Table 1. Data on the large-scale structure, unobtainable with the HMT, were derived from the Berkeley single-dish survey^{5,6} and were added to the synthesis maps. The continuum emission was subtracted from these 'composite' maps to give final 'channel' maps. Photographic representations of some of these 'channel' maps, made on the Cambridge 'Starlink' node, are shown in Fig. 1. These maps will provide detailed information about the ISM in this direction, but here we restrict ourselves to a discussion of the distance to 3C58, and its consequences.

The absorption and column density profiles for H I along the line of sight to 3C58, for each channel, are shown in Fig. 2. Actual on-axis temperatures for the continuum (T_{cont}), and for each channel (T_{chan}) were taken directly from the continuum and 'channel' maps. The emission temperature (T_{em}) in the direction of 3C58, for each channel, was estimated from the surrounding area on each of the 'channel' maps. The uncertainties in these estimates, ranging from 2 to 6 K, are almost solely due to the complicated structure in the ISM (and are taken to represent 2σ errors). From these temperatures the optical depth (τ) and column density ($N_{\text{H I}}$, m⁻²), were derived for each channel using⁷

$$\tau = \log_e(T_{\text{cont}}/(T_{\text{chan}} - T_{\text{em}}))$$

$$N_{\text{H I}} = 1.823 \times 10^{22} \Delta V T_{\text{em}} \text{ m}^{-2} \quad \text{where } \Delta V = 2.0 \text{ km s}^{-1}$$

Figure 2a shows only two distinct absorption features, one from -14.3 to -17.6 km s⁻¹, corresponding to the edge of the local arm, and another centred at -34.1 km s⁻¹, corresponding to the Perseus arm. It is clear, however, from Fig. 2b and the 'channel' maps that in this direction there are two emission peaks in the Perseus arm. The one with lower velocity coincides with the absorption feature at -34.1 km s⁻¹, where a filament crosses the line of sight (Fig. 1). A minimum is found at -39.1 km s⁻¹, beyond which the emission rises to a further peak at -45.7 km s⁻¹. Because no increase in absorption occurs past -39.1 km s⁻¹ corresponding to this second peak, we must conclude that 3C58 is behind the filament at -34.1 km s⁻¹, but in front of the neutral hydrogen at -39.1 km s⁻¹ and beyond. We can thus place 3C58 as being at a distance equivalent to a velocity of -36.6 ± 2.5 km s⁻¹.

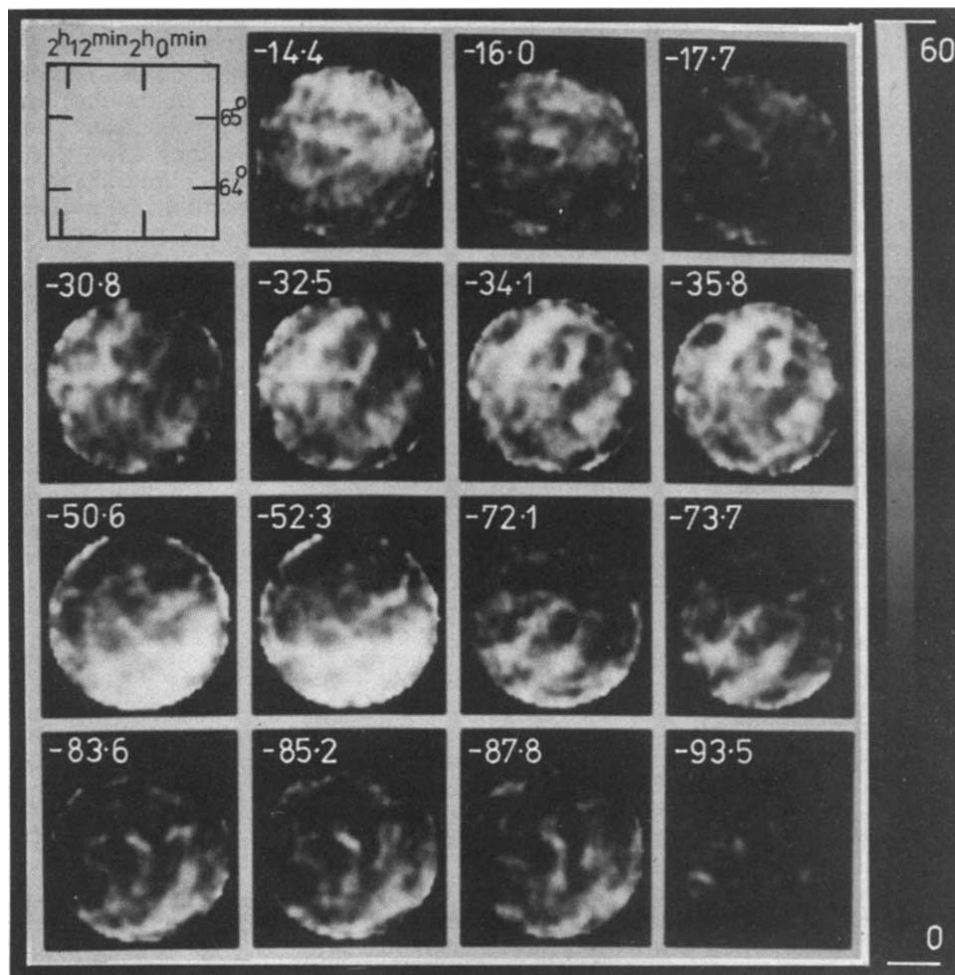
On the basis of the Schmidt rotation model⁸, the distance to 3C58 is then 2.6 ± 0.2 kpc. However, note that any systematic difference from the Schmidt rotation curve would give rise to larger errors. These differences could be due to streaming motions, or any general inadequacy of the Schmidt model. For example, the work of Roberts⁹ would put 3C58 only 2 kpc away, whereas the use of Blitz's rotation curve¹⁰ gives a distance of 3.4 kpc. In the following discussion we have, nevertheless, taken 3C58 to be at a distance of 2.6 ± 0.2 kpc.

Our value for the distance to 3C58 is only a third of the lower limits derived from previous observations, and it is important to understand the reason for such large discrepancy. Consider, for example, Williams' single-dish measurements². He estimated the emission by averaging values at one beamwidth (9 arc min) from the source in four directions. Now it is apparent from Fig. 1 that the ISM shows structure on this scale. If the direction of the source coincides, over a particular velocity range, with a hole in the H I, then the off-axis measurements will give an artificially high value of the emission. This value, combined with the on-axis measurement, will in turn give an artificially high value for the absorption. Bright filaments adjacent to the line of sight will similarly lead to an erroneously high estimate of absorption. This is indeed the case for velocities around -53 , -62 to -75 and -93 km s⁻¹ (Fig. 1), which are just those velocities where Williams finds apparent peaks in absorption. Conversely, a bright filament across the line of sight to the source will mean that the apparent emission and absorption are too low. This again accounts for differences between our results and those of Williams. The effect is particularly noticeable around -85 km s⁻¹ (Fig. 1), where we estimate an emission on the line of sight to 3C58 which is about 10 K higher than that of Williams. This explains why Williams' results shows a drop in absorption to zero in this region, and presumably

Table 1 Specification of the HMT as used for observations in the direction of 3C58

Primary beam	94 arc m in HPBW (field of view shown on Fig. 1 is 80 arc min)
Spatial resolution	7.1×7.9 arc min (RA $\times$ Dec)
Receiver; wide band	10 MHz, phase switched Stokes parameters observed $I + Q$ and $I - Q$
Receiver; narrow band	Digital cross correlation spectrometer measuring $I + Q$ 80 delay channels/spacing 32 frequency channels for each synthesis Bandwidth 0.25 MHz overall
Noise on synthesis maps	~ 0.4 K
Velocity coverage	-14.3 to -108.4 km s ⁻¹ w.r.t. LSR 58 informative channels
Velocity resolution	2 km s^{-1} (to half power points)
Separation of velocity channels	1.65 km s^{-1}
Position of field (195.0.0)	RA = 02 h 02 min Dec = $64^\circ 35'$

Fig. 1 Photographic representations of some of the channel maps. The radio images are positive (bright emission appears white) and cover a range 0–60 K, as shown by the scale on the right. The field of view is sharply cut off at a radius of 80 arc min. The velocity of each channel (w.r.t. the local standard of rest in km s^{-1}) is marked above each map.



enhanced emission was observed. Our results show no significant absorption, nor enhanced emission. The results obtained by Hughes *et al.*³ are based on only one interferometer spacing, and therefore suffer from a lack of information on the large scale structure of the ISM. The interferometric observations of Goss *et al.*⁴ were very poorly sampled both in space and velocity, so detailed comparison with our results is not worthwhile.

3C58, with an assumed distance of around 8 kpc, has been used as a calibrator for several^{11–13} plots of surface brightness (Σ) against linear diameter (D) for SNRs. As most of the calibrators for these plots are shell-types, the relationships derived are representative of idealized shell-type SNRs. There is some theoretical basis¹⁴ for expecting a general evolutionary Σ – D trend for such remnants, although it has been pointed out¹⁵ that young and old remnants should not be considered together. Our revision of the distance to 3C58 will only slightly alter the relations derived, and not at all that of Milne¹⁶, who did not use 3C58 for calibration. On the other hand, the fact that our results, and recent observations^{17,18} of Tycho's SNR (3C10), indicate H I distances much less than was previously thought, must raise doubts concerning other H I distances that have been used for Σ – D calibration. Note^{19,20} that most 'Crab-like' remnants seem to be subluminal compared with an idealized shell-type of the same diameter. Adopting a distance of 2.6 kpc, this is now also true of 3C58, thus strengthening this suggestion.

Our observations give the integrated H I column density between -14.4 km s^{-1} and -36.6 km s^{-1} as $1.0 \pm 0.05 \times 10^{25} \text{ m}^{-2}$. The column density out to -14.4 km s^{-1} can be estimated from Williams' data as $1.55 \pm 0.3 \times 10^{25} \text{ m}^{-2}$, thus the total integrated column density of neutral hydrogen in front of 3C58 is $2.55 \pm 0.3 \times 10^{25} \text{ m}^{-2}$. This is in good agreement with

the value of $2.0 \pm 0.5 \times 10^{25} \text{ m}^{-2}$ recently derived from 'Einstein' X-ray observations²¹.

We can now estimate the optical extinction (A_V) and colour excess ($E(B - V)$) towards 3C58 to deduce its absolute magnitude at maximum ($M_V(\text{max})$). Using

$$N_{\text{tot}} / (6.2 \times 10^{25}) = E(B - V) = A_V / 3.2$$

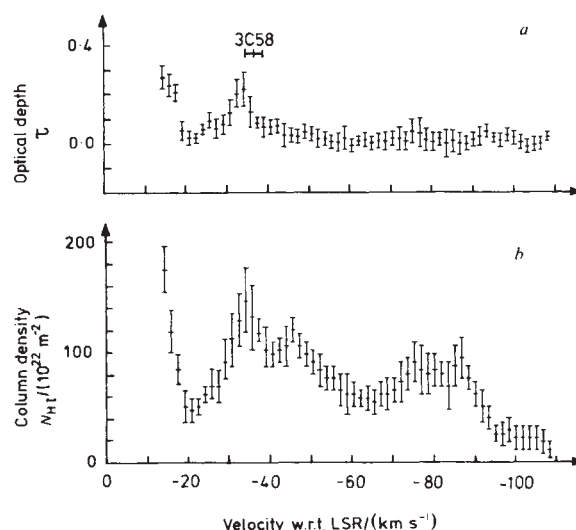


Fig. 2 *a*, Plot of optical depth (τ) against velocity. The velocity of 3C58 is marked. *b*, Plot of column density $N_{\text{H I}}$ against velocity. Error bars indicate estimated 2σ errors on both plots.

Table 2 Comparison of the properties of 3C58 and of the Crab Nebula

	3C58	Crab	Refs
Galactic coordinates	G130.7+3.1	G184.6-5.8	
Distance (kpc)	2.6	2.0	This paper, 29*
Angular size (arc min)	10×6	7×5	29*
Linear size (pc)	7.6×4.5	4×3	
Explosion date	AD 1181	AD 1054	1
Mean expansion velocity (km s ⁻¹)	4,600	2,100	
Distance from galactic plane (pc)	140	200	
Total flux density at 1 GHz (Jy)	33	1,000	29*
Surface brightness at 1 GHz (W m ⁻² Hz ⁻¹ sr ⁻¹)	2.6×10 ⁻¹⁹	1.2×10 ⁻¹⁷	16
Spectral index α ($s \propto \nu^{-\alpha}$)	0.09	0.26	29*
Luminosity (W)			
Radio (10 ⁷ –10 ¹¹ Hz)	1.7×10 ²⁷	1.8×10 ²⁸	29*
X-ray (0.1–4 keV)	1.0×10 ²⁷	2.5×10 ³⁰	21*
Absolute magnitude at maximum (mag)	-13.4	-18	This work, 30

* And refs therein.

(refs 22, 23) we obtain $A_V = 1.3 \pm 0.2$ mag and $E(B-V) = 0.40 \pm 0.05$ mag. For a distance of 2.6 ± 0.2 kpc, the distance modulus is then $(m_V - M_V) = 12.1 \pm 0.1$ mag, where the quoted error does not take account of the possible systematic errors noted above. The apparent magnitude of 3C58 at brightest was¹ $m_V(\text{max}) = 0$ mag so, after allowing for absorption, we obtain: $M_V(\text{max}) = -13.4$ mag. The error in this is dominated by the uncertainty of the apparent magnitude at maximum, which is hard to estimate, but it seems unlikely that this value for the absolute magnitude is in error by more than one magnitude. This makes 3C58 substantially subluminal compared with Type II supernovae observed in external galaxies²⁴, but the detection of extra-galactic supernovae is biased towards those of higher luminosity. Supernovae of Type II have, in any case, a wide range of intrinsic luminosities (as is reasonable if they represent the explosions of very massive stars with extended envelopes).

Table 2 gives values of various properties of 3C58, and those of the Crab Nebula for comparison. It can be seen that 3C58 and the Crab, although similar in many respects, are very different in others. The similarities include their radio morphology—both are ‘filled-centre’ remnants, with filaments seen on high resolution radio maps; their low spectral indices; and their apparent sizes. They also have similar linear sizes, and mean expansion velocities. On the other hand, they seem very different energetically, particularly at other than radio frequencies. The difference between their radio luminosities is only a factor of ~ 10 , whereas for X rays it is $>2,000$. Optically, there are large differences as the filaments of the Crab can be easily studied but the optical emission from 3C58 is barely visible²⁵. Also, the absolute magnitudes of the supernovae that gave rise to these remnants were very different.

There have been several discussions of the properties of ‘Crab-like’ remnants, including attempts to derive some evolutionary trends^{21,26,27} for this class of object, even though so few have been found and well studied. The discussion of Weiler and Panagia²⁶ rests heavily on the observational properties of just three of this class, the Crab, 3C58 and Vela X. They derive an evolutionary sequence, assuming all three to be in the adiabatic or ‘phase II’ expansion²⁸, of the form

$$Sd^2 \propto t^{-0.8}$$

where S is the flux density, d the distance and t the age of the remnant. With the revised distance, it is impossible to fit all three remnants on the same sequence. This could mean that some of the observational data is in error, or the theory may be wrong, or not applicable to all three remnants. If the theory

is substantially correct, and if all three remnants are in ‘phase II’ expansion and are indeed to be placed in the same class, then this class must have a very wide spread of intrinsic properties. On the other hand, Shklovskii²⁷ regards 3C58 as in ‘phase I’, undecelerated, expansion into the ISM. He compares radio-emitting Type II supernovae with ‘filled-centre’ remnants, which are taken to be the products of Type II supernovae. In particular he assumes 3C58 to be a more or less typical ‘filled-centre’ remnant, at a distance of 8 kpc. So the empirical conclusion drawn, that the radio luminosity for Type II supernovae is inversely proportional to age, is no longer valid. The discussion by Becker *et al.*²¹ concerns six filled-centre remnants for which soft X-ray observations are available. They put forward an evolutionary sequence, that with age there is an increase of radio size, and a decrease of both the X-ray luminosity and the ratio of X-ray to radio luminosity. They suggest that the identification of 3C58 with SN1181 is incorrect, and that the probable age is some several thousand years. With a distance of 2.6 kpc for 3C58, this difficulty does not arise.

We thank all those who helped make the observations and reduce the data. D.A.G. thanks the SERC for financial support.

Received 1 June; accepted 30 July 1982.

- Stephenson, F. R. Q. *Jl R. astr. Soc.* **12**, 10–38 (1971).
- Williams, D. R. W. *Astr. Astrophys.* **28**, 309–311 (1973).
- Hughes, M. P., Thompson, A. R. & Colvin, R. S. *Astrophys. J. Suppl.* **23**, 323–370 (1971).
- Goss, W. M., Schwarz, U. J., & Wesselius, P. R. *Astr. Astrophys.* **28**, 305–307 (1973).
- Williams, D. R. W. *Astr. Astrophys. Suppl.* **8**, 505–516 (1973).
- Weaver, H. & Williams, D. R. W. *Astr. Astrophys. Suppl.* **8**, 1–504 (1973).
- Verschuur, G. L. in *Galactic and Extra-galactic Radio Astronomy* (eds Verschuur, G. L. & Kellermann, K. I.) 27–47 (Springer, Berlin, 1974).
- Schmidt, M. *Stars and Stellar Systems* Vol. 5 (eds Blaauw, A. & Schmidt, M.) 513–530 (University of Chicago Press, 1965).
- Roberts, W. W. *Astrophys. J.* **173**, 259–285 (1972).
- Blitz, L. *Astrophys. J. Lett* **231**, L115–L119 (1979).
- Ilovaisky, S. A. & Lequeux, J. *Astr. Astrophys.* **18**, 169–185 (1972).
- Clark, D. H. & Caswell, J. L. *Mon. Not. R. astr. Soc.* **174**, 267–305 (1976).
- Gobel, W., Hirth, W. & Furst, E. *Astr. Astrophys.* **93**, 43–47 (1981).
- Shklovskii, I. S. *Supernovae* (Wiley, New York, 1969).
- Gull, S. F., *Mon. Not. R. astr. Soc.* **161**, 47–69 (1973).
- Milne, D. K. *Aust. J. Phys.* **32**, 83–92 (1979).
- Albinson, J. S. & Gull, S. F. in *Regions of Recent Star Formation* (eds Roger, R. S. & Dewdney, P. E.) 193–199 (Reidel, Dordrecht, 1982).
- Chevalier, R. A., Kirshner, R. P. & Raymond, J. C. *Astrophys. J.* **235**, 186–191 (1980).
- Weiler, K. W. & Shaver, P. A. *Astr. Astrophys.* **70**, 389–397 (1978).
- Lockhart, I. A., Goss, W. M., Caswell, J. L. & McAdam, W. B. *Mon. Not. R. astr. Soc.* **179**, 147–152 (1977).
- Becker, R. H., Helfand, D. J. & Szymkowiak, A. E. *Astrophys. J.* **255**, 557–567 (1982).
- Spitzer, L. *Physical Processes in the Interstellar Medium* (Wiley, New York, 1978).
- Jenkins, E. B. & Savage, B. D. *Astrophys. J.* **187**, 243–255 (1974).
- Tammann, G. A. in *Supernovae* (ed. Schramm, D. N.) 95–116 (Reidel, Dordrecht, 1977).
- van den Bergh, S. *Astrophys. J. Lett* **220**, L9–L10 (1978).
- Weiler, K. W. & Panagia, N. *Astr. Astrophys.* **90**, 269–282 (1980).
- Shklovskii, I. S. *Soviet Astr. Lett.* **7**, 263–264 (1981).
- Woltjer, L. A. *Rev. Astr. Astrophys.* **10**, 129–158 (1972).
- Wilson, A. S. & Weiler, K. W. *Astr. Astrophys.* **49**, 357–374 (1976).
- Chevalier, R. A. in *Supernovae* (ed. Schramm, D. N.) 53–62 (Reidel, Dordrecht, 1977).

Resolution of controversy concerning the morphology of polyacetylene

James C. W. Chien, Y. Yamashita, J. A. Hirsch, J. L. Fan, M. A. Schen & Frank E. Karasz

Department of Polymer Science and Engineering, Materials Research Laboratory, University of Massachusetts, Amherst, Massachusetts 01003, USA

A free standing film of polyacetylene, first obtained by Ito *et al.*¹, has been shown by electron micrograph (EM) to consist of fibrils with average diameter of 20 nm. The results were confirmed by us² and many other laboratories. To eliminate any morphological changes resulting from post polymerization handling of the specimen, we have developed the technique of direct polymerization onto EM grids^{2,3}. The 20–100 nm thick polyacetylene films thus obtained were found to have 3-nm diameter microfibrils as the ultimate morphological entity which aggregate to form the more abundant 20-nm fibrils. Electron

diffraction showed the polymer chain axis to be along the fibre axis³⁻⁵. In marked contrast Wegner and co-workers⁶⁻¹¹ reported that polyacetylene particles contain irregularly shaped lamellae ~5–20 nm in thickness connected into ribbon-like aggregates with polymer chain axis perpendicular to the aggregate direction^{6,7}, and proposed an irregular chain folded lamellar morphology, highly cross-linked with the average length of linear chain segments of ~20 units. We have resolved this controversial observation by reproducing the work of Wegner *et al.* and show here that such specimens were readily transformed to smooth fibrillar morphology by washing with methanolic-HCl.

The direct polymerization of acetylene onto EM grids with Ziegler–Natta catalyst was as described elsewhere^{2,3}. Acetylene was also polymerized by the $\text{Ti}(\text{O}i\text{Bu})_4/4\text{AlEt}_3$ catalyst in the diluent by stirring. The three concentrations of catalyst used were: low $[\text{Ti}] = 1 \text{ mM}$; medium $[\text{Ti}] = 5 \text{ mM}$, and high $[\text{Ti}] = 40 \text{ mM}$. Only for the highest concentration did free standing film form; the others gave suspended particles of polymer. After washing with toluene, methanol or methanolic-HCl, the polymer was collected with carbon coated grids from the suspension to simulate the technique reported by Wegner *et al.*⁶⁻¹¹.

Direct polymerization of acetylene onto gold EM grids with Luttinger catalyst was carried out as follows. A reactor containing suspended EM grids was evacuated for 14 h, flame dried, filled with Ar and transferred into a dry-box. NaBH_4 (20 mg) in 5 ml ether was added to the reactor and then taken out under Ar. The temperature of the reactor was reduced to -78°C and 1 ml of a 10% $\text{Co}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ solution in ethanol was added to the NaBH_4 . The colour immediately turned purple. The reactor was degassed quickly and shaken so that drops of catalyst solution were caught on the grid. Acetylene at ~15 torr was admitted to ensure rapid polymerization. Polyacetylene films formed both on the reactor wall and on the grid. The total time of reaction was 5 min and 60 torr of monomer was consumed. In this procedure the catalyst concentration was $[\text{Co}(\text{NO}_3)_2] = 30 \text{ mM}$ and $[\text{NaBH}_4] = 48 \text{ mM}$.

Using Luttinger catalyst of lower concentrations and polymerizing at -30°C produces only suspended polymer particles. For instance, mixtures of $3.5 \times 10^{-4} \text{ M}$ $\text{Co}(\text{NO}_3)_2$ and $5.3 \times 10^{-3} \text{ M}$ NaBH_4 solutions at -30°C was magenta colour which during acetylene polymerization changed to blue, dark blue, and finally black. In 17 min, the reaction consumed ~50 torr of monomer. Doubling the concentration of both catalyst components caused more rapid polymerization of acetylene at

-30°C . Only suspended polymer particles were produced in these conditions.

Figure 1a shows a transmission EM image of $[\text{CH}]_x$ obtained at a low concentration of $[\text{Ti}] = 1.0 \text{ mM}$. The toluene-washed specimen consisted of fibrils, ribbons and a dark spongy mass; no pseudo-lamellae particles were present. After the polymer was washed with methanolic-HCl, clean fibril morphology resulted (Fig. 1b). The spongy material seems to be soluble in methanolic-HCl and the ribbons in Fig. 1a seem to be aggregates of fibrils glued together by methanolic-HCl soluble substances such as catalyst residue or low molecular weight polymers. At a $[\text{Ti}]$ of 40 mM the $[\text{CH}]_x$ after a thorough toluene wash has a major portion of the specimen masked by spongy substance (Fig. 1c). There are also aggregates of irregularly shaped lamellae-like particles, and fibrils with striation. In fact, this micrograph closely resembles those published by Wegner *et al.* (Fig. 1 of ref. 6, Fig. 3 of ref. 11, Fig. 2 of ref. 8, Fig. 4 of ref. 10, and Fig. 2 of ref. 9.) By washing with methanolic-HCl we obtain EM Fig. 1d which contains only $[\text{CH}]_x$ fibrils although the number of fibrils per unit area in this sample is much greater than that of Fig. 1b. This high $[\text{Ti}]$ is still lower than that commonly used in the preparation of $[\text{CH}]_x$ films, which is 0.2 M. At an intermediate catalyst concentration of 5 mM, the toluene-washed sample has EM Fig. 1e which is intermediate between those of Fig. 1a and c). Thus contribution of the catalyst residue to the dirty morphology seems to be established.

Polyacetylene obtained with Luttinger catalyst $[0.35 \text{ mM Co}(\text{NO}_3)_2/5.3 \text{ mM NaBH}_4]$, washed with ethanol shows definite fibril morphology (Fig. 2a). There were also impurity materials and the fibrils appear to have deposits of foreign materials. Using 10-fold higher catalyst concentration, we observed EMs similar to those reported by Wegner *et al.*⁶⁻¹¹. These are a dense mass fringed with irregularly shaped particles as shown in Fig. 2b. If the polymer was washed with methanolic-HCl, fibrillar morphology is revealed (Fig. 2c). However, much spongy material still remained. It is apparently more difficult to remove contaminants from the polyacetylene obtained with the Luttinger catalyst than in the case of the Ziegler–Natta catalyst.

To show that the two catalyst systems produce truly identical $[\text{CH}]_x$ we have polymerized acetylene directly onto the grid using both the Luttinger (Fig. 3a) and the Ziegler–Natta catalysts (Fig. 3b). The fibrils in the former are smooth, free of foreign substances and have diameters generally greater than in the latter. Electron diffraction patterns of selected areas with parallel bundles of fibrils in both specimens are identical to

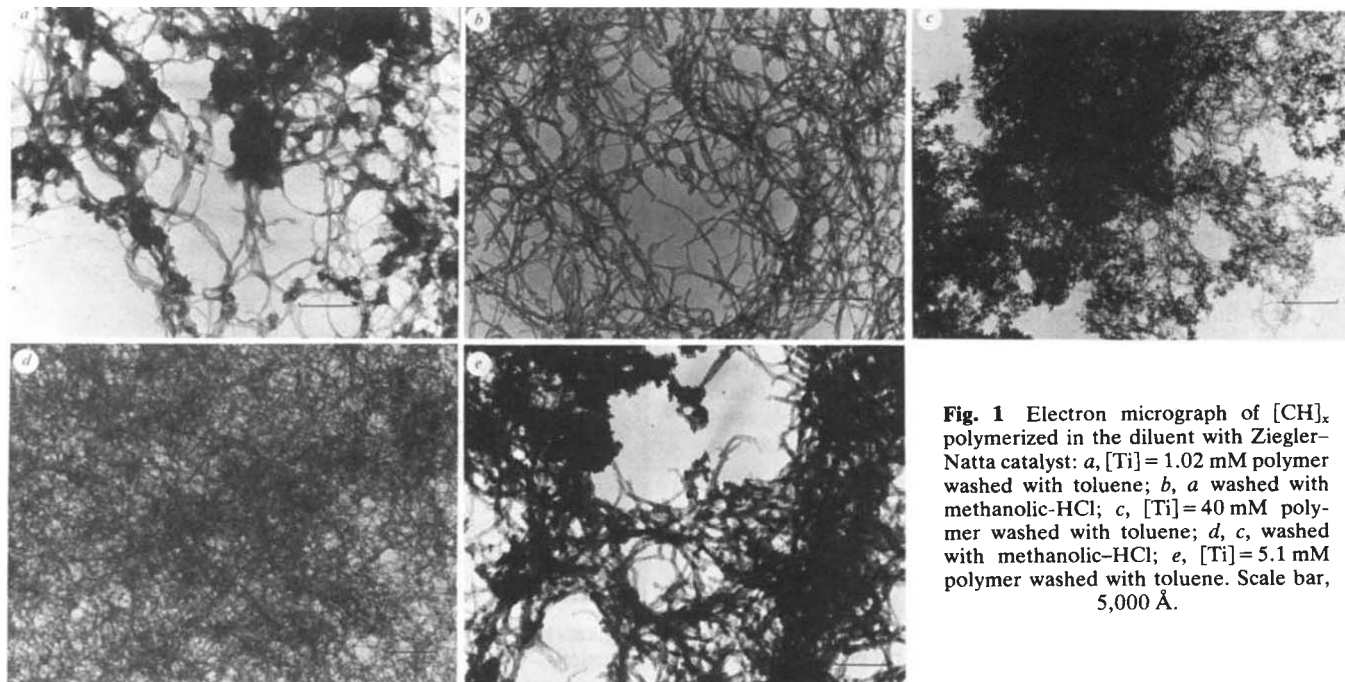


Fig. 1 Electron micrograph of $[\text{CH}]_x$ polymerized in the diluent with Ziegler–Natta catalyst: a, $[\text{Ti}] = 1.02 \text{ mM}$ polymer washed with toluene; b, a washed with methanolic-HCl; c, $[\text{Ti}] = 40 \text{ mM}$ polymer washed with toluene; d, c, washed with methanolic-HCl; e, $[\text{Ti}] = 5.1 \text{ mM}$ polymer washed with toluene. Scale bar, 5,000 Å.

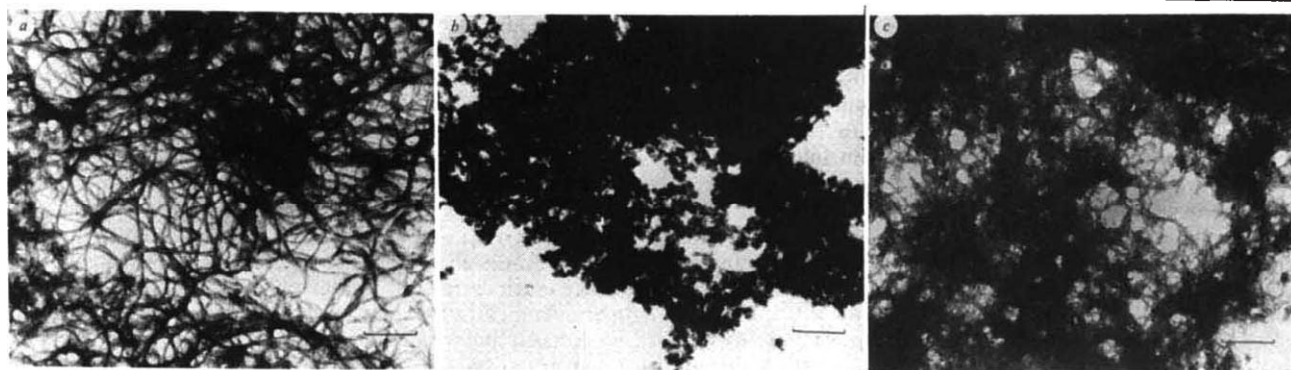


Fig. 2 Electron micrograph of $[\text{CH}]$ -polymerized in the diluent with Luttinger catalyst: *a*, $[\text{Co}] = 0.35 \text{ mM}$, polymer washed with ethanol; *b*, $[\text{Co}] = 30 \text{ mM}$ polymer washed with ethanol; *c*, *b* washed with methanolic- HCl . Scale bar, $5,000 \text{ \AA}$.

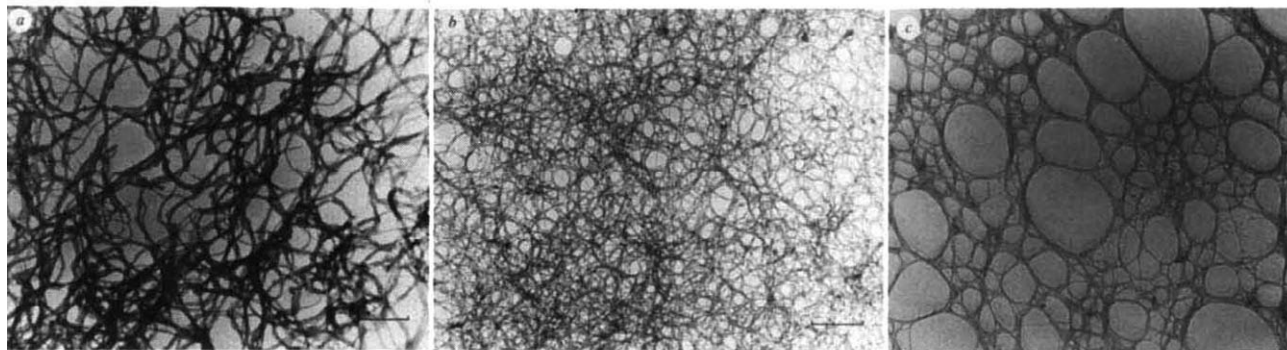


Fig. 3 Electron micrograph of $[\text{CH}]_x$ polymerized directly onto EM grids: *a*, with Luttinger catalyst; *b*, with Ziegler-Natta catalyst; *c*, showing 3-nm diameter microfibrils. Scale bar, $5,000 \text{ \AA}$.

those of *cis*- $[\text{CH}]_x$ described earlier⁵ with the chain axes both along the fibre axis. Furthermore, using low monomer pressure, low Ziegler-Natta catalyst concentration, and short polymerization time, we were able to obtain ultra-thin $[\text{CH}]_x$ specimens which showed an abundance of 3-nm diameter microfibrils (Fig. 3*c*). These microfibrils are probably the ultimate morphological entity which aggregate to form the more commonly reported 20 nm fibrils.

Meyer¹² reported that low-voltage electron-beam damage and OsO_4 staining revealed fibrils of $[\text{CH}]_x$ with regularly spaced striation. It was interpreted to be lamellae packed like stacks of salami slices. We have also occasionally observed such morphology (Fig. 4) in samples that are always aged and sometimes washed with methanolic- HCl . These striations are probably oxidation- or solvent-induced stress cracking. Such cracks appeared as voids along the original fibril and will be more or less regularly spaced because a crack will release the stress in its vicinity. OsO_4 is certainly too strong an oxidizing agent to be used as a staining agent for $[\text{CH}]_x$. Freshly prepared $[\text{CH}]_x$ specimens show neither striation by EM nor long periodicity in low-angle X-ray scattering. Therefore, in the case of Mayer's work the striations were probably produced by stress cracking due to oxidation by OsO_4 or radiation damage.

Wegner *et al.*^{8,10,11} proposed that polyacetylene becomes highly cross-linked even at low temperatures and in the mildest conditions. This was partly based on the observation that the polymer soon becomes difficult to chlorinate to soluble derivatives. It was stated that the average length of linear chain segments is ~ 20 units. ^{13}C -NMR and ozonolysis results showed that there is $<1\%$ of sp^3 carbon atoms. We have found that electron diffraction patterns of pristine *cis*- $[\text{CH}]_x$ has discrete arcs in the equator but diffuse arcs in the layer lines. Careful isomerization in optimal conditions (175°C for 3 min) yielded *trans*- $[\text{CH}]_x$ having very sharp diffraction layer lines suggesting an annealing effect improving the longitudinal order. If *cis*- $[\text{CH}]_x$ were as highly cross-linked as claimed by Wegner *et al.* all of the above observations would seem inconceivable.

The catalyst concentration used in acetylene polymerization is very much higher than in olefin polymerization.^{13,14} When

acetylene is polymerized with a thin coating of catalyst solution on the wall of the reactor or directly onto EM grid, the catalyst residues can be substantially removed by washing with toluene, pentane or heptane. Even then, the Ti content in the free standing polyacetylene is high (0.022%)¹⁵; it can be reduced to 0.008% by methanolic- HCl washing which is a standard work-up procedure of Ziegler-Natta polymerized polyethylene and polypropylene. Washing with toluene or aliphatic solvent would not be enough to remove the catalyst residues. Furthermore, we have found by radiotagging^{16,17} that there are in the diluent very low molecular weight polymers of ~ 500 whereas the molecular weight of the film on the reactor wall is $\sim 22,000$. The dirty morphology of the specimen collected from suspended particle therefore contains copious amounts of catalyst debris

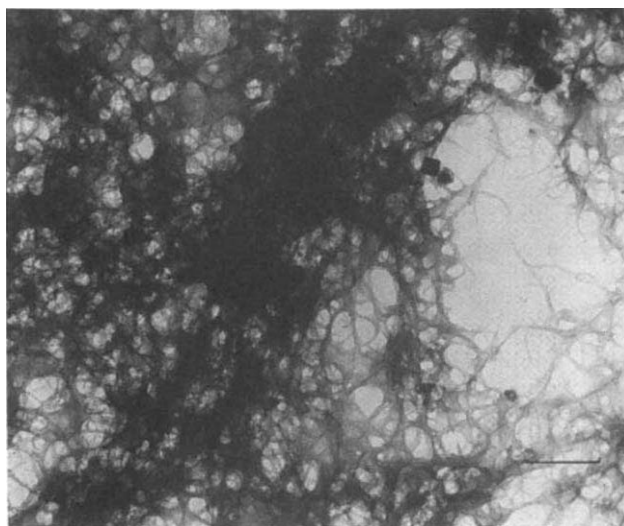


Fig. 4 Electron micrograph of old $[\text{CH}]_x$ specimen showing striations which are probably stress-cracking. Scale bar, $5,000 \text{ \AA}$.

and low molecular weight polymers adhered to the fibrils which can be removed by methanolic-HCl washing.

In conclusion, the basic morphological unit in *cis*-polyacetylene is the 3-nm microfibril which contains about 40 $[\text{CH}]_x$ chains based on its crystal structure^{4,5}. Monomers crystallized from polyacetylene are inserted with the chain axis along the fibre axis. There is no evidence for a chain-folded lamellar morphology which would be difficult for a polymer with a stiff back-bone to assume. Striations of the fibrils sometimes observed are not believed to be basic to the polyacetylene morphology.

Received 15 April; accepted 29 July 1982.

1. Ito, T., Shirakawa, H. & Ikeda, S. *J. Polym. Sci. Polym. Chem. Ed.* **12**, 11 (1974).
2. Karasz, F. E. *et al. Nature* **282**, 286 (1979).
3. Shimamura, K., Karasz, F. E., Hirsch, J. A. & Chien, J. C. W. *Makromolekul. Chem. Rapid Commun.* **2**, 473 (1981).
4. Chien, J. C. W., Karasz, F. E. & Shimamura, K. *Macromolecules* **15**, 1012 (1982).
5. Shimamura, K., Karasz, F. E., Chien, J. C. W. & Hirsch, J. A. *J. Polym. Sci. Polym. Lett. Ed.* **20**, 97 (1982).
6. Lieser, G., Wegner, G., Müller, W. & Enkelmann, V. *Makromolekul. Chem. Rapid Commun.* **1**, 621 (1980).
7. Lieser, G., Wegner, G., Müller, W., Enkelmann, V. & Meyer, W. H. *Makromolekul. Chem. Rapid Commun.* **1**, 627 (1980).
8. Enkelmann, V., Lieser, G., Monkenbusch, M., Müller, W. & Wegner, G. *Molec. Cryst. Liq. Cryst.* **77**, 111 (1981).
9. Lieser, G., Monkenbusch, M., Enkelmann, V. & Wegner, G. *Molec. Cryst. Liq. Cryst.* **77**, 169 (1981).
10. Wegner, G. *Angew. Chem. int. Ed.* **20**, 361 (1981).
11. Enkelmann, V., Müller, W. & Wegner, G. *Syn. Metals* **1**, 185 (1979/80).
12. Meyer, W. H. *Mol. Cryst. Liq. Cryst.* **77**, 1237 (1981).
13. Chien, J. C. W. *J. Polym. Sci. A1*, **1**, 425 (1963).
14. Chien, J. C. W., Wu, J. C. & Kuo, C. I. *J. Polym. Sci. Polym. Chem. Ed.*, (in the press).
15. Chien, J. C. W. *et al. Macromolecules* **15**, 614 (1982).
16. Wnek, G. E. *et al. in Conductive Polymers* (ed. Seymour, R. B.) 183–208 (Plenum, New York, 1981).
17. Chien, J. C. W., Capistran, J. D., Karasz, F. E., Dickinson, L. C. & Schen, M. A. *J. Polym. Sci., Polym. Lett. Ed.* (in the press).

Iron in north-east Pacific waters

R. M. Gordon, J. H. Martin & G. A. Knauer

Moss Landing Marine Laboratories, PO Box 223, Moss Landing, California 95039, USA

Although Fe is an element of great biological¹ and geochemical² importance, little is known about its distribution in the sea. The reasons for this are: (1) contamination is extremely difficult to avoid during sampling and laboratory procedures, not only because of man's wide use of this element, but also because it is fourth most abundant element in the Earth's crust (5.63%)³; (2) the chemistry of Fe is very complex, and its form (or forms) in seawater is poorly known, hence whether one preconcentration technique will work for existing species is questionable. Iron also appears to be very insoluble⁴ in oxygenated ocean water, and most (90%)⁵ precipitates out in association with dissolved organics during estuarine mixing processes^{5–8}. Indeed, some argue that truly dissolved Fe does not exist in seawater and that the fraction found in filtrates is totally colloidal⁹. We have been attempting oceanic dissolved Fe measurements for the past four years and report here three vertical Fe profiles (Fig. 1) that have the following features in common: Fe is severely depleted (0.15–0.30 nmol kg⁻¹) in surface waters; Fe maxima (up to 2.6 nmol kg⁻¹) occur in association with oxygen minima; and, Fe levels appear to vary little in mid-depth waters (0.5–1.0 nmol kg⁻¹).

Our data were obtained using the same sampling (Go-Flo bottles on Kevlar line), filtration (0.4 µm Nuclepore), preconcentration (APDC, DDDC/chloroform organic extraction) and analytical (graphite furnace atomic absorption spectrophotometry, AAS) techniques that were developed largely by Bruland and used by us in previous studies¹⁰. Contamination during collection and filtration procedures was estimated indirectly using Zn and Si analyses; because it is very easy to contaminate samples for Zn (ref. 11), any samples with values

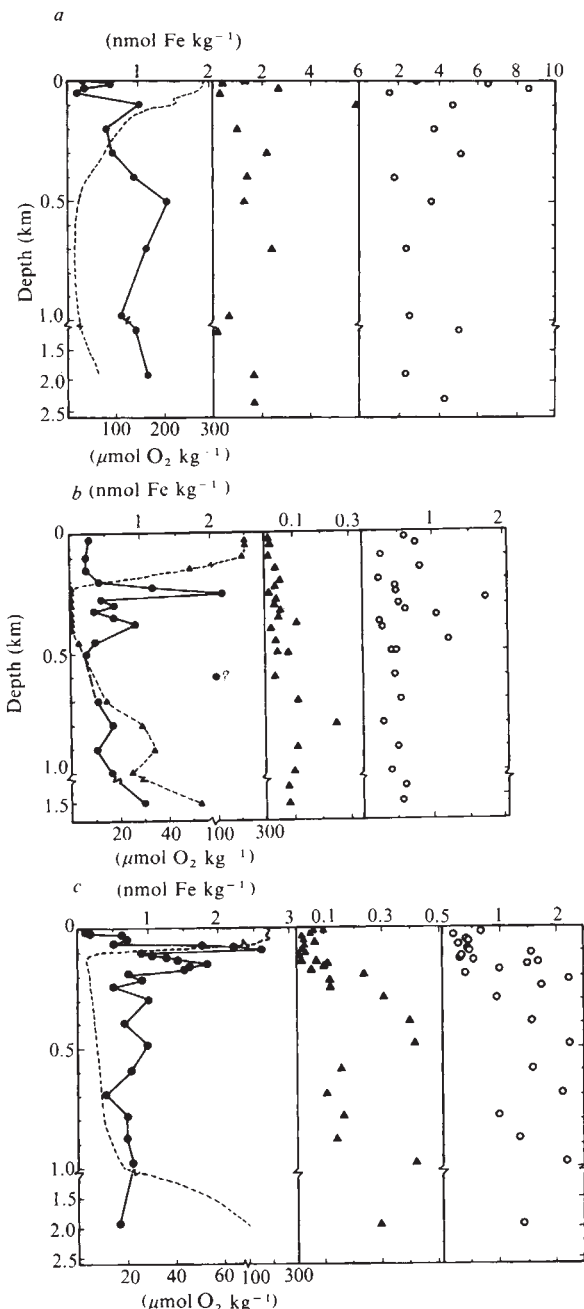


Fig. 1 Dissolved (●), particulate leachable (▲), and particulate refractory (○) Fe versus depth observed: a, off central California (36° 30' N, 123° 05' W) in June 1981; b, hundreds of kilometres west of central Mexico in April 1980; c, 330 km west of Manzanillo, Mexico (18° 00' W, 108° 00' W) in October 1981. (Exact station location of b cannot be given because of contract agreement with Ocean Minerals Corp.)

markedly higher than those expected from Zn/Si ratios (ref. 12) were also assumed to be contaminated with extraneous Fe (Zn/Si data not shown).

Blanks for the organic extraction procedures were ($\bar{x} \pm 1s$) 0.12 ± 0.016 ($n = 7$), 0.13 ($n = 1$) and 0.084 ± 0.023 ($n = 6$) total nmol in the concentrated extract for the data shown in Fig. 1a–c, respectively. Yields for the organic extraction method were estimated in two ways: 95% recovery was obtained using ⁵⁹Fe in the form of Fe³⁺ as a tracer; near 100% recovery was achieved when organic extraction Fe values were compared with those obtained using direct-injection AAS. The latter was possible in low salinity north San Francisco Bay waters where filtrate Fe levels exceeded 0.17 µmol kg⁻¹ (see Table 1).

Table 1 Filtrate Fe levels in north San Francisco Bay waters

Salinity (%)		Direct injection	Organic extraction	Organic extraction
		($\mu\text{mol kg}^{-1}$)	($\mu\text{mol kg}^{-1}$)	Direct injection $\times 100$ (%)
2.2	A	0.286	0.252	88
	B	0.176	0.179	102
6.1	A	14.4	13.8	96
	B	8.50	8.15	96
12.3	A	6.36	7.02	110
	B	6.50	7.36	113
				$\bar{X} \pm 1s$ 101 $\pm$ 9.4

Two replicates were analysed using the organic extraction preconcentration technique; precision was generally good, as exemplified by the data presented in Table 2. Iron values estimated using another preconcentration technique (chelex ion exchange at pH 6) are also shown in Table 2. Good agreement between the two preconcentration methods was generally found in the 11 replicates that were checked. This was also true for the Fig. 1a data (chelex values not shown).

Although there are few recent Fe data available for comparison with those we report here, those that do exist are in general agreement. Landing and Bruland¹³ also reported low Fe levels (0.4 nmol kg^{-1}) in surface waters, and maxima in the oxygen minimum ($1\text{--}2 \text{ nmol kg}^{-1}$) at locations in the same vicinity as our Fig. 1a station. Although we failed to find a

Table 2 Dissolved Fe values found off central Mexico ($18^\circ 00' \text{ N}$, $108^\circ 00' \text{ W}$) in October 1981 (data are also shown in Fig. 1c)

Depth (m)	OE 1	OE 2	$\bar{X}$ OE Fe (nmol kg^{-1})	CX Fe
Surface	0.14	0.13	0.14	0.16
5	0.28*	0.13	0.13	—
10	(3.6)†	(3.8)	(3.7)	—
20	0.16	0.49*	0.16	0.16
30	0.69	0.60	0.64	0.51
40	(1.4)	(1.1)	(1.2)	(1.6)
50	0.73	0.73	0.73	0.61
60	0.49	0.60	0.54	0.64
70	1.6	2.0	1.8	1.7
80	2.2	2.3	2.2	—
90	2.7	2.6	2.6	—
100	0.95	0.89	0.92	—
110	1.2	0.98	1.1	—
120	1.2	1.3	1.3	—
130	1.5	1.5	1.5	—
140	(4.8)	(4.0)	(4.4)	—
150	1.9	1.8	1.8	—
160	1.7	1.4	1.6	—
175	1.7	1.4	1.6	—
195	0.74	0.73	0.74	—
220	1.0	0.82	0.91	—
245	0.66	0.37	0.52	—
295	1.1	0.95	1.0	1.0
395	0.76	0.60	0.68	—
490	0.90	1.0	0.95	—
590	0.92	0.61	0.76	0.86
690	0.50	0.29	0.40	—
785	0.75	0.68	0.72	—
880	0.76	0.65	0.70	0.69
980	0.84	0.76	0.80	—
1,955	0.73	0.46	0.60	0.74

Iron was preconcentrated in 2 replicates (OE 1, OE 2) using organic extraction. Replicate values were averaged ($\bar{X}$ OE Fe) and are compared with those obtained using chelex ion exchange (CX Fe) preconcentration at selected depths.

* Not used in $\bar{X}$.

† () thought to be contaminated (see text).

mid-depth (750–1,500 m) maximum that they reported, we may have missed this feature, as we only had two samples in this depth range.

Low surface Fe values could be expected for two reasons: (1) Fe is least soluble in oxygen-rich surface waters⁴; and (2) phytoplankton require this nanonutrient^{14,15} and depletion is logical in the euphotic zone. Higher Fe levels in association with oxygen minima are also likely because of increased solubility⁴ in such waters. In this regard, it is interesting that the shape of the profiles shown in Fig. 1b and c are the same as those observed in the Black Sea¹⁶, although the Pacific Fe levels are two orders of magnitude lower.

Why Fe varies little in mid-depth waters is uncertain. Perhaps the lack of variability suggests equilibrium with particulate phases, but our particulate Fe data (Fig. 1) generally give little information in this regard. Two fractions of particulate Fe were analysed: refractory Fe (ref. 17) and weakly leachable (25% acetic acid) Fe (ref. 18). The latter technique is not very suitable for this element, as hydrous oxides are not removed. Obviously, more would be learned if a more appropriate procedure were used to separate hydrous Fe oxide coatings from crystalline Fe (ref. 19). Nevertheless, the particulate Fe data at least illustrate the general insolubility of Fe, as these values usually exceeded those for dissolved Fe. Refractory Fe amounts reflect various levels of suspended aluminosilicates in the water column, and as expected, quantities were higher at the near-shore California location (Fig. 1a) than those in the more oceanic waters (Fig. 1b, c). Leachable Fe quantities showed a consistent trend in only one case; values were lowest where dissolved Fe levels were highest at the top of the oxygen minimum off Mexico (Fig. 1c). The increase in leachable Fe beneath the dissolved Fe maxima suggests that Fe was being removed back onto particles at these depths.

Based on recent experience with several trace elements, it is probable that the $\mu\text{g per kg}$ amounts of dissolved Fe previously reported for seawater resulted from gross sample contamination. The data we present here and those of Landing and Bruland¹³ are an order of magnitude less than the maximum possible dissolved amounts (20 nmol kg^{-1}) estimated by Byrne and Kester²⁰. Thus, it is possible that the Fe we measured was truly dissolved, perhaps in organic complexes^{5–8} or as $\text{Fe}(\text{OH})_3$ (refs 4, 20). Obviously, much remains to be learned about Fe in oceanic waters, and we hope that this report will stimulate research on this topic.

We thank W. W. Broenkow for the CTOD oxygen data, and S. E. Fitzwater for help with sample collection and analyses. This research was supported by grants from the NSF Marine Chemistry (VERTEX, OCE 80-03200) and Biological Oceanography (OCE 79-09431) programmes, the Ocean Minerals Corporation, and the US Environmental Protection Agency (CR 8077110020).

Received 7 April; accepted 29 July 1982.

1. Bowen, H. J. M. *Trace Elements in Biochemistry* (Academic, London, 1966).
2. Goldberg, E. D. *J. Geol.* **62**, 249–265 (1954).
3. Taylor, S. R. *Geochim. cosmochim. Acta* **28**, 1273–1285 (1964).
4. Stumm, W. & Morgan, J. J. *Aquatic Chemistry* 2nd edn (Wiley, New York, 1981).
5. Boyle, E. A., Edmond, J. M. & Sholkovitz, E. R. *Geochim. cosmochim. Acta* **41**, 1313–1324 (1977).
6. Sholkovitz, E. R., Boyle, E. A. & Price, N. B. *Earth planet. Sci. Lett.* **40**, 130–136 (1978).
7. Sholkovitz, E. R. *Earth planet. Sci. Lett.* **41**, 77–86 (1978).
8. Sholkovitz, E. R. & Copland, D. *Geochim. cosmochim. Acta* **45**, 181–189 (1981).
9. Mill, A. J. B. *Envir. Tech. Lett.* **1**, 97–108 (1980).
10. Bruland, K. W., Franks, R. P., Knauer, G. A. & Martin, J. H. *Analyt. chim. Acta* **105**, 233–245 (1979).
11. Bruland, K. W., Knauer, G. A. & Martin, J. H. *Nature* **271**, 741–743 (1978).
12. Bruland, K. W. *Earth planet. Sci. Lett.* **47**, 176–198 (1980).
13. Landing, W. M. & Bruland, K. W. *EOS* **62**, 906 (1981).
14. Lewin, J. C. & Chen, C. H. *Limnol. Oceanogr.* **16**, 670–675 (1971).
15. Menzel, D. W. & Ryther, J. H. *Deep-Sea Res.* **7**, 276–281 (1961).
16. Spencer, D. W. & Brewer, P. G. *J. geophys. Res.* **76**, 5877–5892 (1971).
17. Eggemann, D. W. & Betzer, P. R. *Analyt. Chem.* **48**, 866–890 (1976).
18. Chester, R. & Hughes, M. J. *Chem. Geol.* **2**, 249–262 (1967).
19. Gibbs, R. J. *Bull. geol. Soc. Am.* **88**, 829–843 (1977).
20. Byrne, R. H. & Kester, D. R. *Mar. Chem.* **4**, 255–274 (1976).

Radiocaesium and plutonium in intertidal sediments from southern Scotland

A. B. MacKenzie & R. D. Scott

Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

The dispersion of radionuclides discharged to the Irish Sea from the works of British Nuclear Fuels Ltd at Sellafield (Windscale) is a topic of major importance, attracting considerable interest in recent years. Annual surveys by the Ministry of Agriculture, Fisheries and Food (MAFF) of marine radioactivity^{1,2} supplemented by other work³⁻¹⁷ have defined the major features of this dispersion during the 30 yr of the plant's operation. Two species of radiological concern are radiocaesium isotopes (¹³⁷Cs, β , γ , $t_{1/2} = 30.25$ yr; ¹³⁴Cs, β , γ , $t_{1/2} = 2.05$ yr) which constitute the main effluent radioactivity and plutonium isotopes (²³⁸Pu, α , $t_{1/2} = 86$ yr; ²³⁹Pu, α , $t_{1/2} = 2.44 \times 10^4$ yr; ²⁴⁰Pu, α , $t_{1/2} = 6.58 \times 10^3$ yr) with high radio-toxicities. We present here plutonium and radiocaesium profiles for intertidal sediment cores collected from four sites in southern Scotland (Fig. 1) during October 1979 and discuss the present and longer-term significance of the interaction of Sellafield effluent with such sediments.

Intertidal sediment cores of length 18 cm were extruded immediately on collection and divided into 3-cm sections. The sediments consisted of sand/silt and, with the exception of Westferry, shell fragments. The biological population was restricted to small crustaceans and worms at Westferry but was more varied at the other sites. Dried sediment samples were analysed for radiocaesium¹⁸, plutonium^{19,20}, aluminium, manganese and carbonate, results being presented in Table 1.

Plutonium and radiocaesium arise in the study area from two main sources—Sellafield and nuclear weapons testing fallout. Sellafield effluent has had a ¹³⁴Cs/¹³⁷Cs ratio in the range 0.1–0.25 in recent years⁸ and a ²³⁸Pu/^{239,240}Pu ratio increasing from 0.05 to 0.25 during the period 1966–79^{1,2,6}. Fallout contains ¹³⁷Cs but not ¹³⁴Cs and has a ²³⁸Pu/^{239,240}Pu ratio of about 0.05 (ref. 21). The Sellafield discharge is very much larger than

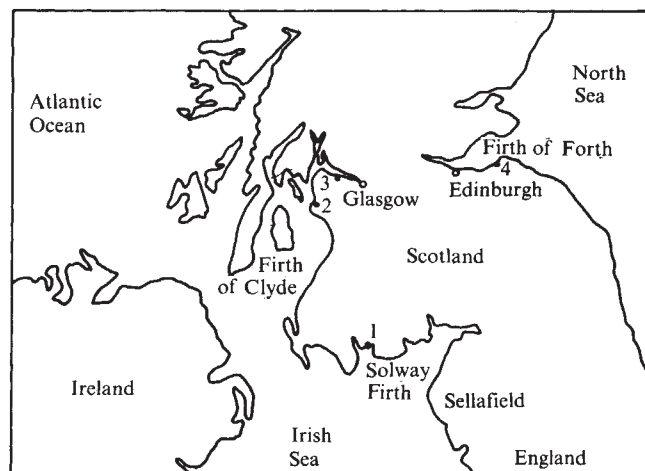


Fig. 1 Map showing locations of sampling sites. 1, Skyreburn Bay; 2, Ardrossan; 3, Westferry; 4, Aberlady Bay.

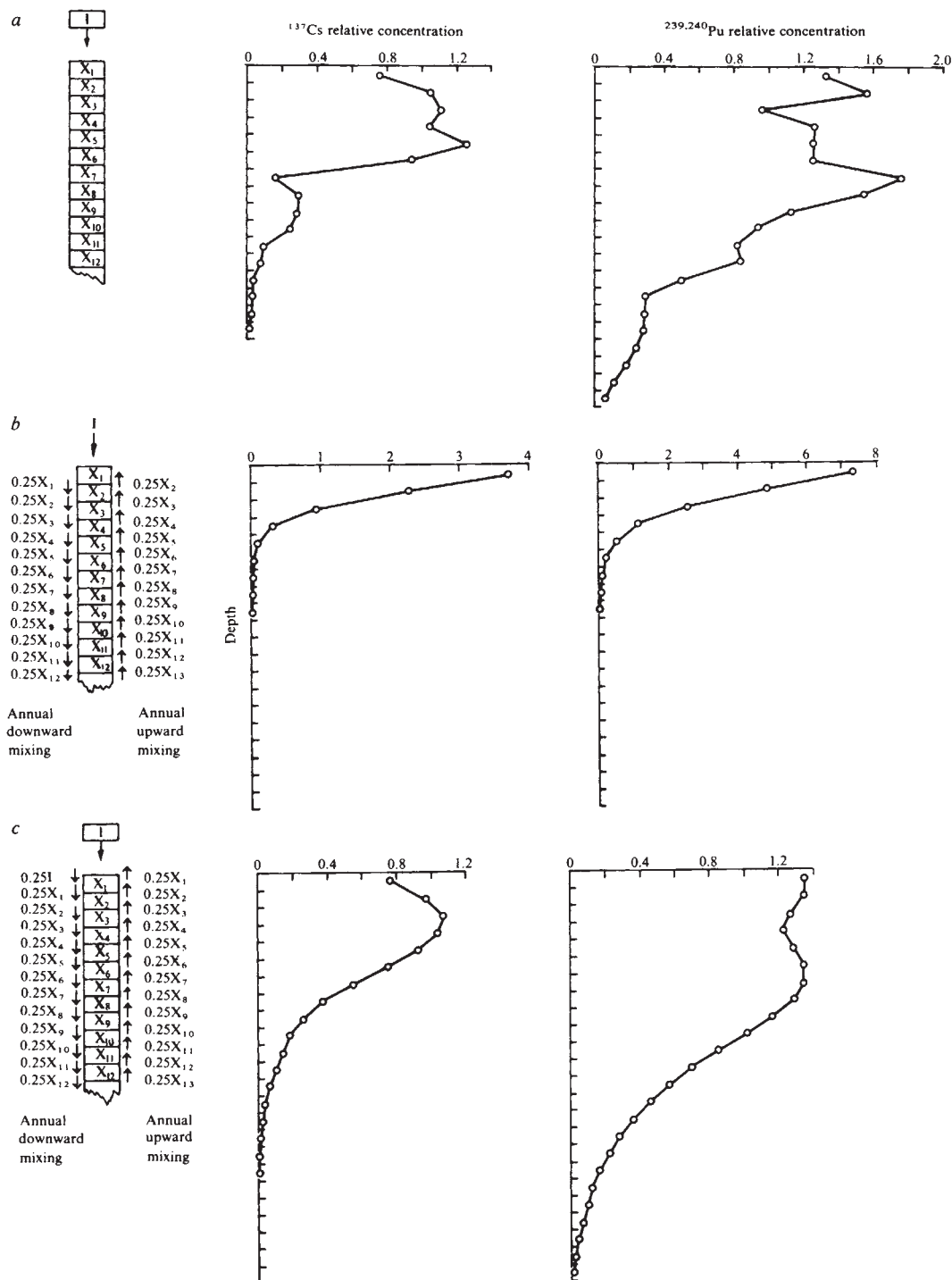
that from any other nuclear facility in the UK and concentrations of manmade radionuclides in the marine environment of south-west Scotland are consequently dominated by Sellafield waste^{1,2,8}. The east coast site will receive a minor input from Dounreay (and possibly La Hague) but the largest source of manmade radionuclides in this section of the North Sea is Sellafield⁷. Consistent with this, it is apparent from Table 1 and the isotope ratios given above that all four sites studied are contaminated to some extent with Sellafield effluent. The Skyreburn Bay concentrations can be almost totally attributed to Sellafield and are in good agreement with MAFF data for nearby sites^{1,2}. The other sites show ¹³⁷Cs and ^{239,240}Pu contents which are factors of 10 and 10², respectively, lower than those in Skyreburn Bay, reflecting dilution and preferential uptake of Pu by sediments during transit. As the concentrations at the three northerly sites are similar to soil fallout levels, the observed isotope ratios could result from mixtures of either Sellafield effluent and fallout, or wastes discharged from Sellafield at different times. However, the results imply a lower limit of 70–75% for the Sellafield component of the total activity. In attempting to rationalize the observed profiles, it is

Table 1 Analytical results

Location	Depth (cm)	¹³⁷ Cs		^{239,240} Pu		Al concentration (% dry weight)	Mn concentration (% dry weight)	CaCO ₃ concentration (% dry weight)
		concentration (mBq g ⁻¹)	¹³⁴ Cs/ ¹³⁷ Cs	concentration (mBq g ⁻¹)	²³⁸ Pu/ ^{239,240} Pu			
Skyreburn Bay	0–3	478 ± 5	0.073 ± 0.004	58 ± 2	0.23 ± 0.01	3.8	0.076	4.4
	3–6	657 ± 5	0.072 ± 0.003	43 ± 2	0.22 ± 0.01	3.5	0.051	4.7
	6–9	762 ± 5	0.062 ± 0.003	49 ± 2	0.23 ± 0.01	3.7	0.047	4.6
	9–12	950 ± 7	0.059 ± 0.002	77 ± 3	0.22 ± 0.01	3.5	0.046	5.3
	12–15	793 ± 6	0.060 ± 0.003	63 ± 3	0.21 ± 0.01	3.4	0.045	5.0
	15–18	742 ± 5	0.050 ± 0.003	42 ± 2	0.25 ± 0.02	3.8	0.045	5.6
Ardrossan	0–3	30 ± 2	0.15 ± 0.08	1.05 ± 0.04	0.15 ± 0.02	1.4	0.011	0.6
	3–6	37 ± 2	0.14 ± 0.06	0.31 ± 0.02	0.15 ± 0.02	1.5	0.011	2.3
	6–9	45 ± 2	0.08 ± 0.05	0.68 ± 0.03	0.15 ± 0.02	2.1	0.012	2.7
	9–12	48 ± 2	0.06 ± 0.05	0.83 ± 0.03	0.14 ± 0.02	1.7	0.013	3.3
	12–15	55 ± 2	0.08 ± 0.05	1.12 ± 0.04	0.13 ± 0.02	1.8	0.013	5.0
	15–18	55 ± 2	BDL	0.98 ± 0.04	0.15 ± 0.02	1.8	0.014	7.4
Westferry	0–3	87 ± 3	0.09 ± 0.02	1.22 ± 0.03	0.18 ± 0.02	3.3	0.026	<0.4
	3–6	53 ± 2	0.09 ± 0.03	0.07 ± 0.003	0.20 ± 0.02	4.5	0.026	<0.4
	6–9	25 ± 2	BDL	0.05 ± 0.003	0.16 ± 0.02	4.1	0.028	<0.4
	9–12	12 ± 1	BDL	0.021 ± 0.004	BDL	3.7	0.033	<0.4
	12–15	9 ± 1	BDL	0.015 ± 0.001	BDL	3.3	0.032	<0.4
	15–18	3 ± 1	BDL	0.005 ± 0.002	BDL	4.2	0.023	<0.4
Aberlady Bay	0–3	39 ± 2	0.06 ± 0.03	0.53 ± 0.02	0.13 ± 0.01	2.6	0.017	3.1
	3–6	26 ± 2	0.08 ± 0.03	0.34 ± 0.02	0.20 ± 0.02	2.3	0.016	2.2
	6–9	15 ± 1	BDL	0.24 ± 0.010	0.15 ± 0.02	2.1	0.014	0.5
	9–12	11 ± 1	BDL	0.11 ± 0.005	0.19 ± 0.02	1.5	0.008	3.6
	12–15	11 ± 2	BDL	0.075 ± 0.008	BDL	1.6	0.011	2.5
	15–18	13 ± 2	BDL	0.046 ± 0.003	0.13 ± 0.02	1.7	0.011	4.2

Radionuclide errors are based on 1 σ counting statistics; errors on other results ~3%. Al and Mn were determined by instrumental neutron activation analysis; CO₃ were determined by measurement of CO₂ evolved by reaction with acid. BDL, below detection limit.

Fig. 2 Sediment box models and resultant radionuclide profiles relating to the Sellafield discharge of ^{137}Cs (discharge data used for 1964 to 1979 inclusive) and $^{239,240}\text{Pu}$ (discharge data used for 1960 to 1979 inclusive). Relative concentrations can be related between models for one nuclide but cannot be related between ^{137}Cs and $^{239,240}\text{Pu}$. An arbitrary, but constant, depth scale is used in all cases.



Mixing between boxes occurs and in the example selected, intense mixing involving exchange of 0.25 of the material originally present is assumed to occur annually between contiguous boxes. The amount of exchange between boxes can be varied to simulate different intensities of mixing and varied rates of mixing can also be applied down the sediment column. **c**, 'Mixing-plus-accumulation' model. This model is a direct combination of models **a** and **b**. (More complex models of this type are discussed by Santschi *et al.*²³.)

necessary to consider the well defined^{8,10,12} temporal variations in the Sellafield discharge in conjunction with processes which influence the uptake and subsequent behaviour of radionuclides in sediment. Surface sediments are subject to a complex interaction of influences, including physical (for example, mixing, compaction, variable accumulation/erosion), chemical (such as redox reactions, interstitial water processes, sediment heterogeneity) and biological (mixing, irrigation, feeding, excreting) processes²², resulting in severe difficulties in modelling the system. Figure 2 shows three models combining the Sellafield discharge data with varied accumulation and mixing situations. These are examples from a suite of models

and illustrate that 'accumulation-plus-mixing' profiles are similar to those for 'accumulation-only' but with the curve smoothed, variations attenuated and radionuclide penetration to greater depth. Subsurface maxima are retained, even with intense mixing, provided there is also accumulation, but the non-symmetrical maximum for ^{137}Cs can be displaced towards the surface relative to that for $^{239,240}\text{Pu}$.

The Skyreburn Bay and Ardrossan profiles (Fig. 3) conform to the 'accumulation-only' or 'accumulation-plus-mixing' type whereas Westferry and Aberlady Bay relate to the 'mixing-only' model. No correlation is apparent between radionuclide profiles and those for Al (indication of clay content), Mn (indication of

redox processes) and carbonate (shell content), showing that these parameters do not exert a major influence on the radionuclide distributions and that variations in input, sedimentation and mixing are probably the dominant processes. Sellafield radionuclides dispersed by water movements require a transit time of 8 months or less to reach the northern parts of the Clyde Sea area^{8,24}, so a shorter transit time to Skyreburn Bay and Ardrossan can be assumed, as is confirmed by the profiles for these sites, which show trends relating to the Sellafield discharge up to 1979. Sediment mixing by waves, currents and biological activity was observed at both locations so the profiles should be related to the 'accumulation-plus-mixing' model. Relating maxima in the profiles to those in the Sellafield discharge implies 'sedimentation rates' in the range 2.5–3.0 cm yr⁻¹ for these sites. This 'sedimentation' is probably a transient component of a complex, long-term reworking of these sediments because the bays have been stable over long periods^{25–27}. An extended study of Skyreburn Bay, to be reported elsewhere, has shown that cores collected from different parts of the bay show similar profiles, the total depth of radionuclide penetration is ~35 cm and variations in the profiles as a function of time tend to support the above hypothesis of reworking. Distinction between reworking and continuous sedimentation is important because the former would result in retention of the radionuclides in the surface zone of potential radiological significance over much longer timespans.

Several workers have studied sedimentation processes in the River Esk, Ravensglass Estuary, by relating radionuclide concentrations and isotope ratio profiles to the Sellafield discharge^{6,10,12,14,17}, and Clifton and Hamilton present a detailed study of the influence of geochemical parameters, variable sedimentation rates and mixing processes on radionuclide profiles and derived chronologies. Aston and Stanners¹² contend that their ²³⁸Pu/^{239,240}Pu profiles for Ravensglass Estuary sediments, in conjunction with a model deriving a vanishingly small 'apparent diffusion coefficient', demonstrate negligible post-depositional remobilization of plutonium. However, the information which can be derived on this subject either from the present work or from that of Aston and Stanners is subject to considerable uncertainty because: (1) Sellafield discharge ²³⁶Pu/^{239,240}Pu data are not available before 1978; the only data available for comparison with sediment profiles are Hetherington's results for surface sediments from the Ravensglass Estuary for the period 1966–75 and these results have an implicit sampling error of several per cent and an explicit 1σ error on counting statistics of as much as 47%. (2) Mixing of discharges of different ages and isotope ratios in Irish Sea waters and surface sediments leads to uncertainty in relating observed profiles to discharge data. (3) Experimental errors for ²³⁸Pu/^{239,240}Pu ratios are often as large as 10%. (4) Sampling increments are large by comparison with any plausible remobilization movement of plutonium. Thus, while this type of study does provide useful information on the uptake of radionuclides by sediments, it does not provide a sensitive method for study of possible remobilization and removal processes which would be more meaningfully studied by analysis of radionuclide concentrations in interstitial and overlying waters^{6,28}.

Westferry and Aberlady Bay occupy relatively sheltered sites (in contrast to Skyreburn Bay and Ardrossan) and show radionuclide profiles which suggest negligible accumulation with mixing which results in penetration of ¹³⁷Cs and ^{239,240}Pu to depths of at least 18 cm over a period of about 30 yr. ¹³⁷Cs and ^{239,240}Pu have similar distributions in Aberlady Bay, indicating that similar processes influence both species. For Westferry, however, there is pronounced penetration of ¹³⁷Cs to greater depths than ^{239,240}Pu. Livingston and Bowen²⁹ suggest that similar observations for nearshore sediments can be explained by diagenetic remobilization of Pu and this argument could be invoked for the Westferry profiles. However, alternative explanations also exist, for example, preferential downward migration of radiocaesium in interstitial waters, the ambiguity again

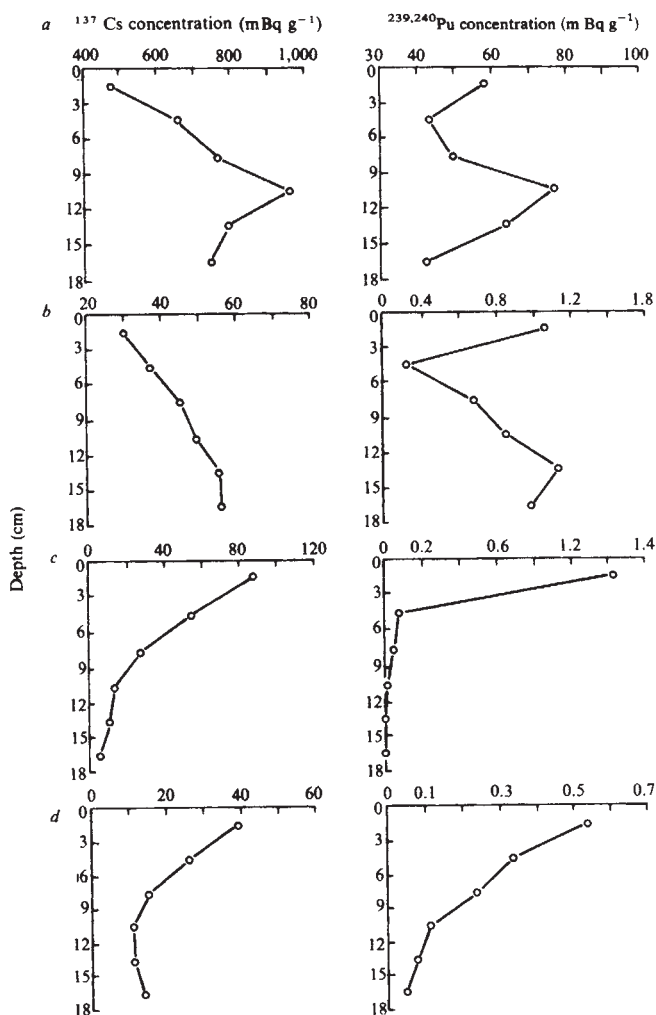


Fig. 3 ¹³⁷Cs and ^{239,240}Pu profiles in intertidal sediments from southern Scotland, October 1979. a, Skyreburn Bay; b, Ardrossan; c, Westferry; d, Aberlady Bay.

illustrating the limitations of bulk sediment analysis in investigating such processes.

Uncertainties concerning possible removal impose a limitation on estimates of long-term variations of radionuclide concentrations in these sediments. However, assuming no removal occurs other than by radioactive decay and that radionuclide concentrations increase in response to a continued discharge from Sellafield at present levels, an (improbably long) operation period of 200 yr would result in ¹³⁷Cs and ^{239,240}Pu inventories increasing by factors of 3 and 10, respectively, relative to 1979 levels. Resultant concentrations at the study sites would then, at highest, equal present levels in Cumbria, the significance of which has been discussed by Hunt^{1,2}. This calculation derives an order of magnitude upper limit and the concentrations which develop will almost certainly be lower than this, as removal by processes other than radioactive decay will occur and the Sellafield discharge will probably decrease as the nuclear power industry evolves and improved chemical procedures (for example, the SIXEP ion exchange plant³⁰) are introduced.

Partial financial support of this work by the International Atomic Energy Agency under research contract no. 2409/RB is gratefully acknowledged. We thank Professor H. W. Wilson for his encouragement.

Received 8 March; accepted 3 August 1982.

1. Hunt G. J. *Aquatic Environment Monitoring Report No. 4* (Ministry of Agriculture Fisheries and Food, Lowestoft, 1980).
2. Hunt G. J. *Aquatic Environment Monitoring Report No. 6* (Ministry of Agriculture Fisheries and Food, Lowestoft, 1981).
3. Jefferies, D. F., Preston, A. & Steele, A. K. *Mar. Pollut. Bull.* 4, 118 (1973).

4. Wilson T. S. R. *Nature* **248**, 125 (1974).
5. Livingston H. D. & Bowen, V. T. *Nature* **269**, 586 (1977).
6. Hetherington, J. A. *Mar. Sci. Commun.* **4**, 239 (1978).
7. Murray, C. N., Kautsky, H., Hoppenheit, M. & Domian M. *Nature* **276**, 225 (1978).
8. Baxter, M. S., McKinley, I. G., MacKenzie, A. B. & Jack, W. *Mar. Pollut. Bull.* **10**, 116 (1979).
9. Murray, C. N., Kautsky, H. & Eicke, H. F. *Nature* **278**, 617 (1979).
10. Aston, S. R. & Stanners, D. A. *Estuar. coastal mar. Sci.* **9**, 529 (1979).
11. McKinley, I. G., Baxter, M. S., Ellett, D. J. & Jack, W. *Estuar. coastal Shelf Sci.* **13**, 69 (1981).
12. Aston, S. R. & Stanners, D. A. *Nature* **289**, 581 (1981).
13. Day, J. P. & Cross, J. E. *Nature* **292**, 43 (1981).
14. Stanners, D. A. & Aston, S. R. *Estuar. coastal Shelf Sci.* **13**, 101 (1981).
15. Stanners, D. A. & Aston, S. R. *Estuar. coastal Shelf Sci.* **13**, 409 (1981).
16. Aston, S. R. & Stanners, D. A. *Estuar. coastal Shelf Sci.* **14**, 167 (1982).
17. Clifton, R. J. & Hamilton, E. I. *Estuar. coastal Shelf Sci.* **14**, 433 (1982).
18. MacKenzie, A. B., Baxter, M. S., McKinley, I. G., Swan, D. S. & Jack, W. *J. radioanalyt. Chem.* **48**, 29 (1979).
19. Coleman, G. R. *The Radiochemistry of Plutonium* (National Academy of Sciences Report NAS-NS 3058, 1965).
20. MacKenzie, A. B. *Analyst. Proc.* (in the press).
21. Cambray, R. S. & Eakins, J. D. *Studies of Environmental Radioactivity in Cumbria Pt 1* (UKAEA report AERE-R-9807, HMSO, London, 1980).
22. Berner, R. A. *Early Diagenesis* (Princeton University Press, 1980).
23. Santschi, P. H., Li, Y. H., Bell, J. J., Trier, R. W. & Kawtaluk, K. *Earth planet. Sci. Lett.* **51**, 248 (1980).
24. McKinley, I. G., Baxter, M. S. & Jack, W. *Estuar. coastal Shelf Sci.* **13**, 59 (1981).
25. *Ordnance Survey Map of Scotland*, Popular Edn, 1" to 1 mile, 3rd revision, 1921-23 (HMSO, London, 1923).
26. *The Atlas of Scotland* (Thomson, Edinburgh, 1832).
27. Ainslie J. *Map Published According to Act of Parliament with Improvements till 1800* (Brown, Edinburgh, 1800).
28. Nelson, D. M. & Lovett, M. B. in *Proc. IAEA conf.*, Vienna 1980 (International Atomic Energy Agency, Vienna, 1981).
29. Livingston, H. D. & Bowen, V. T. *Earth planet. Sci. Lett.* **43**, 29 (1979).
30. British Nuclear Fuels Ltd. *Windscale and Calder Works* (BNFL, Risley, 1981).

Detection of imogolite in soils using solid state ^{29}Si NMR

P. F. Barron*, M. A. Wilson†, A. S. Campbell‡ & R. L. Frost§

* Brisbane NMR Centre, Griffith University, Nathan, Queensland, Australia 4111

† CSIRO Division of Fossil Fuels, PO Box 136, North Ryde, New South Wales, Australia 2113

‡ Department of Soil Science, Lincoln College, Canterbury, New Zealand

§ Queensland Institute of Technology, Brisbane, Australia 4001

High resolution solid state ^{29}Si -NMR has recently been used to examine silicates, and synthetic and natural zeolites¹⁻⁴. This has been possible through the use of dipolar decoupling, magic-angle spinning and cross-polarization techniques which have enabled line narrowing and also considerable signal enhancement for ^{29}Si nuclei in close proximity to protons. It has been found that isotropic ^{29}Si chemical shifts in solids and solutions are generally similar and depend mainly on the degree of condensation of silicon-oxygen tetrahedra¹. For silicates, increasing condensation from single (Q^0 , -66 to -72 p.p.m.) to double tetrahedra (Q^1 , -78 to -82 p.p.m.), to chains (Q^2 , -86 to -88 p.p.m.) and cyclic layered structures (Q^3 , -107 to -109 p.p.m.) leads to increasing diamagnetic shielding. In solid aluminosilicates, the ^{29}Si chemical shift is sensitive to the substitution of aluminium. A further important point was that the ^{29}Si linewidth in aluminosilicates is sensitive to the regularity of aluminium distribution in the lattice. Thus ^{29}Si -NMR has obvious potential for the elucidation of the coordination of silicon in soils and mineral components. We report here the first ^{29}Si -NMR study of the clay mineral imogolite, and show that ^{29}Si -NMR is an excellent analytical technique for the identification of imogolite (or imogolite like materials, that is proto-imogolite) in clay fractions. The observed chemical shift of imogolite (-78 p.p.m.) is consistent with the proposal that imogolite is a tubular hydrated aluminosilicate in which silicon tetrahedra are isolated by coordination through oxygen with three aluminium atoms and one proton.

X-ray diffraction is useful for identifying long range order in materials; that is, order brought about by alignment of large

numbers of atoms such as that existing in the crystalline state. NMR, however, gives information about localized bonding. Largest effects are observed from changes in structure only one or two atoms removed from the nucleus under investigation. The application of ^{29}Si -NMR to the study of mineral and clay chemistry would be expected to have greatest potential in elucidating the structure of those components for which X-ray diffraction gives little structural information, such as the fine clay fractions from volcanic ash soils.

These materials may contain quantities of imogolite⁵⁻¹¹, a hydrated aluminosilicate of tubular structure in which it has been proposed that silicon tetrahedra are isolated by coordination through oxygen with three aluminium atoms and one proton. The external surface of the tube is made up of a gibbsite-like structure, but the internal surface has exposed Si-OH groups which arise from O_3SiOH tetrahedra. Imogolite has been identified by X-ray and electron diffraction patterns, but this requires the presence of bundles of aligned tubes. IR spectroscopy has been used to identify and crudely quantify the presence of imogolite⁵. However, the method is far from satisfactory due to possibility of overlapping bands from other mineral species. There are at present no satisfactory methods for identifying and quantifying the amounts of imogolite in clay fractions.

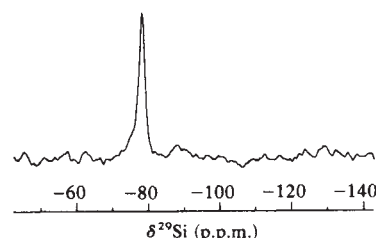


Fig. 1 DD/MAS ^{29}Si -NMR spectrum of imogolite. The chemical shift of imogolite is -78 p.p.m. from TMS.

Previously^{12,13} we reported the characterization of poorly ordered aluminosilicates in a vitric andosol and a spodosol from New Zealand. Imogolite was not identified in any of these samples by electron microscopy, although this may be the result of insufficient resolution. Nevertheless, the samples meet all the defined requirements of allophane¹⁴, and are therefore very suitable for testing the ^{29}Si -NMR analytical method. In the present work the ^{29}Si -NMR spectra of the andosol and spodosol fine clay fractions are investigated and comparison is made with ^{29}Si -NMR spectra of authentic imogolite.

The soils used (Taupo and Bealey) and soil sites have been described elsewhere¹²⁻¹⁵. Chemical properties of the size fractions investigated are shown in Table 1. Authentic imogolite was provided by CSIRO Division of Soils and originated from Kitakami City, Suata Prefecture, Japan and was separated from Kitakami Pumice. ^{29}Si -NMR spectra were obtained at 59.61 MHz on a Bruker CXP-300 spectrometer using cross-polarization with magic angle spinning (CP/MAS) or using a single pulse sequence with dipolar ^1H decoupling (DD/MAS) and B_1 fields of 10 and 50 mT for ^1H and ^{29}Si respectively. The latter experiment can be considered as giving quantitative assessment of the types of ^{29}Si environments present whilst the former experiment should yield information on the presence of hydroxy groups or bound water in close proximity to ^{29}Si nuclei. All the samples were packed into Andrews-type Delrin rotors and spun at speeds of 3-4 kHz. A pulse delay of 20 s was used for DD/MAS experiments. Experiments using a range of pulse delays were performed and showed that delays of 10 s or more would allow quantitative interpretation of DD/MAS spectra. CP/MAS experiments were performed using single 5-ms contacts repeated at 3-s intervals. Contact time dependent experiments were performed for selected samples. Chemical shifts are referenced indirectly to TMS.

The DD/MAS spectrum of the authentic imogolite is shown in Fig. 1. It consists of one peak at -78 p.p.m. According to

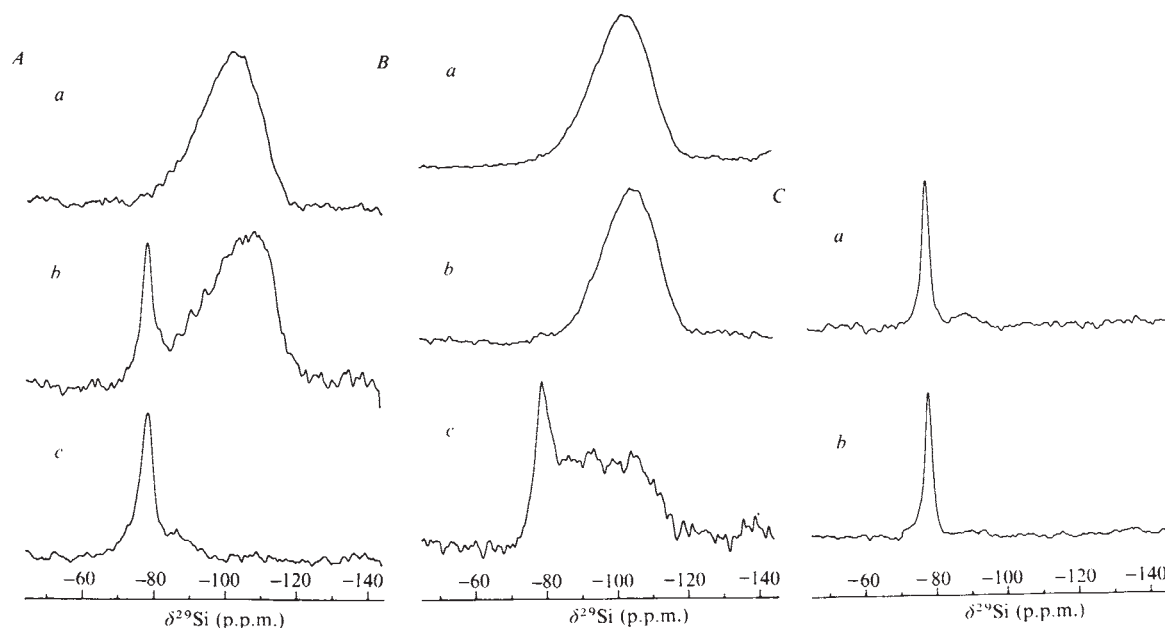


Fig. 2 A, DD/MAS ^{29}Si -NMR spectra of Taupo B horizon fractions. a, 2–150 μm ; b, 0.2–2 μm ; c, 0.2 μm . B, DD/MAS ^{29}Si -NMR spectra of Taupo C horizon fractions. a, 2–150 μm ; b, 0.2–2 μm ; c, <0.2 μm . C, DD/MAS ^{29}Si -NMR spectra of Bealey fractions. a, 0.2–2 μm ; b, <0.2 μm .

the accepted view, imogolite is a $Q^3(3\text{Al})$ aluminosilicate as it consists of trisubstituted silicon tetrahedra. The only published spectra of Q^3 aluminosilicates are those of pyrophyllite ($Q^3(0\text{Al})$) and muscovite ($Q^3(1\text{Al})$) and it is difficult to predict the chemical shift of imogolite from these samples alone. However, Lippmaa¹ has published spectra of a complete range of Q^4 aluminosilicates. The Q^4 samples substituted with three alumina, $Q^4(3\text{Al})$, have chemical shifts at –88 p.p.m. The corresponding Q^4 samples substituted with four silica units and no alumina groups $Q^4(0\text{Al})$ (silica) have a chemical shift (δ) at –107 p.p.m. Thus, the difference in chemical shift between the two samples (Δ) is

$$\begin{aligned}\Delta &= \delta Q^4(0\text{Al}) - \delta Q^4(3\text{Al}) \\ &= (-107) - (-88) = -19 \text{ p.p.m.}\end{aligned}$$

The chemical shift of $Q^3(3\text{Al})$ groups in imogolite can be predicted to a first approximation as

$$\delta Q^3(3\text{Al}) = \delta Q^3(0\text{Al}) - \Delta$$

Chemical shifts of –91 and –95 p.p.m. are observed for $Q^3(0\text{Al})$ tetrahedra in kaolin¹⁶ and pyrophyllite^{1,16} (the second resonance at –91 p.p.m. reported by Lippmaa¹ must be due to the presence of kaolin in the sample studied) respectively. Hence, the expected chemical shift of imogolite is ~–72 to –76 p.p.m. There is thus good agreement between the observed and predicted chemical shift. Our results are therefore in good agreement with the proposed structure of imogolite as a $Q^3(3\text{Al})$ aluminosilicate. That imogolite is a hydroxy aluminosilicate is supported by the signal enhancement obtained using CP/MAS. Contact time dependence experiments showed

maximum enhancement occurring at ~0.75 ms. In contrast, maximum enhancement for kaolin is achieved in the vicinity of 5 ms (ref. 17). The much shorter T_{CH} in imogolite is easily rationalized when the structures are considered. Kaolin contains hydroxyl groups bonded to aluminium in aluminium sheets which are in turn attached to silicon tetrahedral sheets. However, imogolite contains hydroxyl groups directly bound to silicon. Thus, the internuclear ^{29}Si – ^1H separation to nearest protons is probably smaller in imogolite than in kaolin and resulting in more rapid cross-polarization.

DD/MAS ^{29}Si spectra of Taupo B and C horizons, and Bealey soil fractions are shown in Fig. 2 respectively. Each sample was examined by both DD/MAS and CP/MAS techniques. The DD/MAS spectra (Fig. 2A) of the three Taupo B fractions clearly show the variation in the silicate composition with fraction size. The non-clay fraction spectrum exhibits a broad resonance from ~–80 to –110 p.p.m. and centred at ~–100 p.p.m. The unresolved nature of the resonance indicates the presence of an amorphous array of ^{29}Si environments with the shift range, according to Lippmaa's work¹, indicative of both silicates and aluminosilicates. The $\text{SiO}_2/\text{Al}_2\text{O}_3$ ratio of 8.9 previously determined for this fraction indicates that this broad resonance results largely from silicates. Only minor amounts of quartz and cristobalite were previously detected by X-ray analysis and this is consistent with the minor amount of signal in the vicinity of –107 p.p.m. CP/MAS examination of this sample yields very little signal, indicating that silicates containing hydroxy groups or constituent water are not present in significant amounts. The coarse clay fraction spectrum exhibits quite different features. The spectrum shows a resolved

Table 1 Chemical properties of the clay-sized separates of Bealey spodosol and Taupo vitric andosol

	Spodosol			Andosol				
				B horizon		C horizon		
Size fraction (μm)	<0.2	0.2–2	<0.2	0.2–2	2–150	<0.2	0.2–2	2–150
Organic carbon (%)	15.3	14.2	13.2	9.4	0.27	6.3	0.94	0.16
Total Si	13.9	16.7	17.5	26.4	29.8	21.6	31.8	34.7
Al	32.4	28.1	25.4	14.7	6.4	21.7	7.6	6.8
Fe	1.2	1.9	7.9	5.8	2.2	6.0	2.3	2.2
$\text{SiO}_2/\text{Al}_2\text{O}_3$ mole ratio	0.82	1.1	1.3	3.5	8.9	1.9	8.0	9.9

See refs 13, 14 for further details.

signal at -78 p.p.m., consistent with the presence of imogolite-like material as well as a broad signal extending from ~ -90 to -115 p.p.m. and centred at ~ -105 p.p.m.

The spectrum of the Taupo B fine clay fraction virtually consists only of the relatively sharp resonance at -78 p.p.m. indicating the high regularity of the material. Only a small amount of unresolved signal is present over the -80 to -110 p.p.m. range. The previously determined $\text{SiO}_2/\text{Al}_2\text{O}_3$ ratio is further reduced to 1.3. A cross-polarization study of this sample was performed which revealed that this signal cross-polarized and achieved maximum enhancement at ~ 0.75 ms. These findings are also consistent with the proposed structure and NMR results for imogolite. However, the linewidth of the -78 p.p.m. resonance is larger than that for imogolite, probably indicating a lesser degree of order of structure.

Comparison of spectra of Taupo C horizon fractions (Fig. 2B) demonstrates a similar but not identical variation in silicon environments. The non-clay fraction for Taupo C yields a ^{29}Si spectrum identical to that of the B horizon in that it consists of a single broad featureless peak centred at ~ -101 p.p.m. consistent with an amorphous silicate/aluminosilicate composition. However, in contrast to the B horizon coarse clay fraction, the C horizon coarse clay fraction spectrum is identical to that of the non-clay fraction, except for a small shift to ~ -105 p.p.m., and this is consistent with its high $\text{SiO}_2/\text{Al}_2\text{O}_3$ ratio of 8.0. There is no sign of highly ordered and regular material as found in the B horizon. For the fine clay fraction of the C horizon, however, the ^{29}Si NMR spectrum does indicate the presence of ordered material. As for the coarse clay B fraction a clearly resolved resonance is present at -78 p.p.m. with a broad resonance extending from -80 to -110 p.p.m. This broad resonance appears to exhibit partially resolved resonances at ~ -92 and -105 p.p.m. Again, this sharper signal exhibited rapid cross-polarization. This fraction has a considerably lower $\text{SiO}_2/\text{Al}_2\text{O}_3$ ratio (1.9) which is again consistent with the proposal that the -78 p.p.m. resonance is due to imogolite.

Only two fractions from the Bealey soil, the coarse ($0.2-2 \mu\text{m}$) and fine ($<0.2 \mu\text{m}$) clay fractions, were examined (Fig. 2C). Both gave what could be considered as virtually single line spectra with a sharp resonance at -78 p.p.m. The resonance is in fact considerably sharper than that found in the Taupo fractions with a linewidth close to that obtained for imogolite. The previously determined $\text{SiO}_2/\text{Al}_2\text{O}_3$ ratio for the coarse and fine clays of 1.1 and 0.82 respectively are even lower than that found for the B horizon fine clay. Hence, the silicon species in the Bealey clay fractions consist almost entirely of imogolite or perhaps proto-imogolite. Note that at present we cannot distinguish different physical forms of imogolite.

We have been able to detect quantitatively imogolite in clay fractions when it is present in significant amounts, using ^{29}Si DD/MAS NMR. However, the clay fractions constitute only a few per cent¹² of the total soils and hence imogolite is present at levels of less than a few per cent in the whole soils. Quantitative detection of imogolite in the whole soils using ^{29}Si DD/MAS NMR may be difficult, if not impossible, given the time required to obtain spectra of sufficient signal to noise, and dynamic range and chemical shift overlap problems associated with the expected signal from the large excess of amorphous silicates/aluminosilicates. However, the difference in cross-polarization behaviour between imogolite and other aluminosilicates pointed out earlier, and the non-cross-polarizability of silicates may enable non-qualitative detection of imogolite at low levels in whole soils by ^{29}Si CP/MAS NMR using a suitably short contact time. To demonstrate this, we have been able to detect the presence of imogolite in the whole Taupo B horizon soil using the ^{29}Si CP/MAS technique with a contact time of 1 ms. We are currently investigating the detection limits of this technique.

We conclude therefore that:

(1) ^{29}Si -NMR provides a viable method for the detection of imogolite in soils. It is straightforward to quantify the fraction

of silicon present as ordered imogolite, when present in significant amounts, relative to that present as amorphous silicates/aluminosilicates. Due to the cross-polarization characteristics of imogolite, the non-quantitative detection of the presence of imogolite in small amounts relative to silicates and other aluminosilicates may be feasible.

(2) Imogolite is an important component of several clay fractions of an andosol and spodosol from New Zealand. In some cases, it seems to be the only component containing silicon.

(3) The observed chemical shift of imogolite (-78 p.p.m.) and its cross-polarization behaviour is consistent with the proposal that imogolite consists of isolated silica tetrahedra substituted with three alumina and one hydroxyl unit.

Since submission of this manuscript, it has been pointed out that another $\text{Q}^3(3\text{Al})$ aluminosilicate, margarite, has been examined by ^{29}Si -NMR (ref. 17). A single peak at -75.5 p.p.m. was obtained in good agreement with our observations for imogolite.

We thank J. Skjemstad and Mr J. Pickering, CSIRO Division of Soils, for the imogolite sample and M. A. Young for the preparation of some of the clay samples. The Bruker CXP-300 NMR spectrometer is operated by the Brisbane NMR Centre.

Received 4 May; accepted 28 July 1982.

- Lippmaa, E., Magi, M., Samoson, A., Engelhardt, G. & Grimer, A. R. *J. Am. chem. Soc.* **102**, 4889-4893 (1980).
- Lippmaa, E., Magi, M., Samoson, A. & Engelhardt, G. *J. Am. chem. Soc.* **103**, 4992-4996 (1981).
- Klinowski, J., Thomas, J. M., Fyfe, C. A. & Hartman, J. S., *J. phys. Chem.* **85**, 2590-2594 (1981).
- Engelhardt, G., Lippmaa, E. & Magi, M. *JCS chem. Commun.* 712-713 (1981).
- Farmer, V. C., Fraser, A. R., Russell, J. D. & Yoshinaga, N. *Clay Minerals* **12**, 55-57 (1977).
- Cradwick, P. D. G. *et al. Nature phys. Sci.* **240**, 187-189 (1972).
- Farmer, V. C. & Fraser, A. R. in *International Clay Conference 1978* (eds Mortland, M. M. & Farmer, V.) (Elsevier, Amsterdam, 1979).
- Parfitt, R. L. in *Soils with Variable Charge*, 167-179 (N.Z. Soil Bureau, Wellington, New Zealand, 1981).
- Wada, K. & Yoshinaga, N. *Am. Miner.* **54**, 50-71 (1969).
- Wada, K. & Matsubara, I. *Trans. 9th int. Cong. Soil Sci.* **3**, 123-131 (1968).
- Yoshinaga, N. *Soil Sci. Pl. Nutr., Tokyo* **8**, 22-29 (1962).
- Campbell, A. S., Young, A. W., Livingstone, L. G., Wilson, M. A. & Walker, T. W. *Soil Sci.* **123**, 362-368 (1977).
- Young, A. W., Campbell, A. S. & Walker, T. W. *Nature* **284**, 46-48 (1980).
- van Olphen, H. *Science* **177**, 91-97 (1971).
- N.Z. Soil Bureau Bull.* **26**, 1-127 (1978).
- Barron, P. F., Skjemstad, J. & Frost, R. L. *Nature* (submitted).
- Magi, M., Samoson, A., Tarmak, M., Englehardt, G. & Lippmaa, E. *Doklady Akad. Nauk. SSSR* **261**, 1169 (1981).

Reduction of molybdate by soil organic matter: EPR evidence for formation of both Mo(V) and Mo(III)

B. A. Goodman & M. V. Cheshire

The Macaulay Institute for Soil Research, Craigiebuckler, Aberdeen AB9 2QJ, UK

The reducing ability of organic fractions extracted from the soil is well known¹⁻³. Here the reduction of Mo(VI) by humic acid and polysaccharide components from peat and soil, respectively, has been studied by electron paramagnetic resonance (EPR) spectroscopy. We report that reduction to Mo(V) species occurred initially, but with the humic acid, we observed reduction of part of the molybdenum to Mo(III). This represents the first definite evidence for Mo(III) formation in natural substances.

The humic acid⁴ was extracted from a raised bog peat overlying parent materials of the Hatton Association⁵. The polysaccharide was extracted with 0.2 M NaOH from a soil of the Countesswells series, Countesswells Association, and isolated by using Polyclar and ethanol precipitation⁶. For the formation of molybdenum complexes, the humic acid (~ 100 mg) was dissolved in 0.2 M NaOH and mixed with sodium molybdate

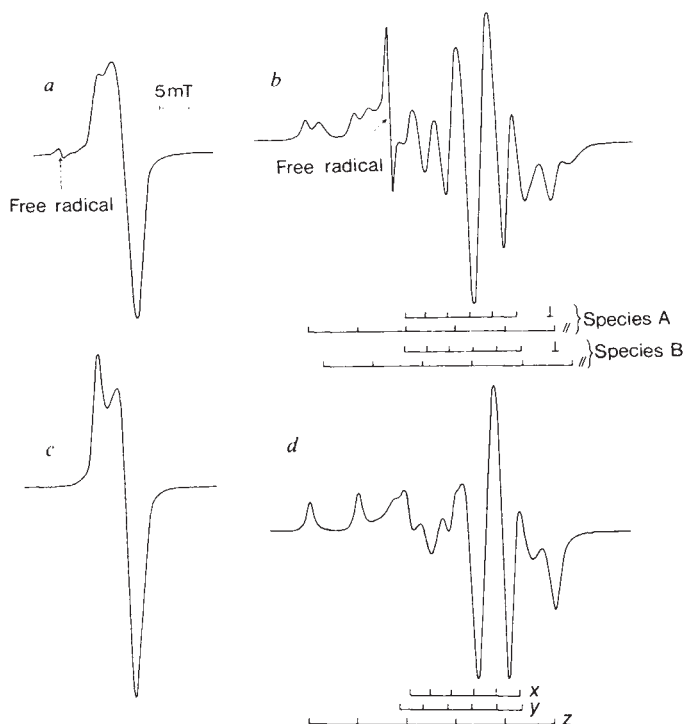


Fig. 1 EPR spectra at ambient temperatures of *a*, the Mo(V)-humic acid complex, with natural isotopic ratios for Mo; *b*, the ^{95}Mo (V)-humic acid complex; *c*, the Mo(V)-polysaccharide complex with natural isotopic ratios for Mo; and *d*, the ^{95}Mo (V)-polysaccharide complex.

solution (1.7 mg MoO_3 in 0.2 M NaOH). After 30 min the solution was acidified to pH 2 with 6 M HCl and centrifuged. The precipitated humic acid-molybdenum complex was dried in air under reduced pressure. The soil polysaccharide (~100 mg) was dissolved in 10 ml water and mixed with sodium molybdate solution (as above). After 30 min the pH was reduced to 3 with HCl and the mixture was freeze-dried. In a subsequent experiment the solid humic acid-molybdenum complex was washed twice with 0.1 M HCl (50 ml) and dried in air under reduced pressure.

EPR spectra were recorded at ambient temperatures at X-band frequencies on a Varian E 104 spectrometer. In the case of humic acid and molybdenum of natural isotopic abundance, a spectrum was obtained (Fig. 1*a*) which had parameters characteristic of axially symmetric Mo(V) species⁷. A weak peak from the well known humic acid free radical was also observed. The molybdenum peaks in Fig. 1*a* arise from the isotopes with zero spin. Weak hyperfine structure from the ^{95}Mo ($I = 5/2$, 15.7%) and ^{97}Mo ($I = 5/2$, 9.5%) isotopes was also observed (not shown), indicating that more than one type of Mo(V) species was present. Therefore, the experiment was repeated using molybdenum enriched to 97.43% in ^{95}Mo (obtained from AERE, Harwell). This produced a spectrum (Fig. 1*b*) in which two components (A and B) are clearly resolved. The spectral parameters for both species, obtained by computer simulation, are given in Table 1. When the same experiment was performed with the polysaccharide, only one species was obtained (Fig. 1*c*, *d*). In this case the absence of a second component in the spectrum allowed the computer simulation to be performed with greater precision than for the humic acid complexes. The peak positions could not be matched accurately when an axially-symmetric Hamiltonian was used, and the parameters given in Table 1 correspond to the best fit obtained when three independent axes were assumed. The components from the humic acid system were fitted by an axially-symmetric Hamiltonian but, because of appreciable line overlap in the perpendicular region, it is not possible to state with certainty that true axial symmetry

Table 1 EPR parameters for molybdenum complexes with soil organic matter

	$g_{\perp}$	$g_{\parallel}$	$A_{\perp}$ (mT)	$A_{\parallel}$ (mT)
Mo(V)-humic acid species A	1.941	1.968	3.6	8.0
Mo(V)-humic acid	1.939	1.953	3.8	8.1
Mo(V) polysaccharide	g_x 1.948	g_y 1.951	g_z 1.979	A_x 3.6
			A_y 4.1	A_z 8.0
Mo(III)-humic acid	g 1.944		A 4.6	

applies. Nevertheless, it is clear from the magnitudes of the g -values that both of the Mo(V)-humic acid complexes differ from the polysaccharide complex.

In an attempt to alter the relative amounts of the Mo(V)-humic acid complexes, the solid molybdenum-humic acid samples were washed with 0.1 M HCl. The resulting spectra from the natural isotope and ^{95}Mo system are shown in Fig. 2*a* and *b*, respectively. These spectra are dominated by the intense narrow component from the humic acid free radical. The hyperfine structure on the low-field side of the spectra arises from copper porphyrin complexes, which are known to be present in the humic acid⁸. The molybdenum components are on the high-field side of the spectra. The natural isotope system produced a single isotropic peak, which was confirmed as arising from molybdenum by the presence of the hyperfine structure in the spectrum obtained for the sample containing the ^{95}Mo isotope. Of the various molybdenum ions that might be expected to give EPR signals at ambient temperatures, only high-spin

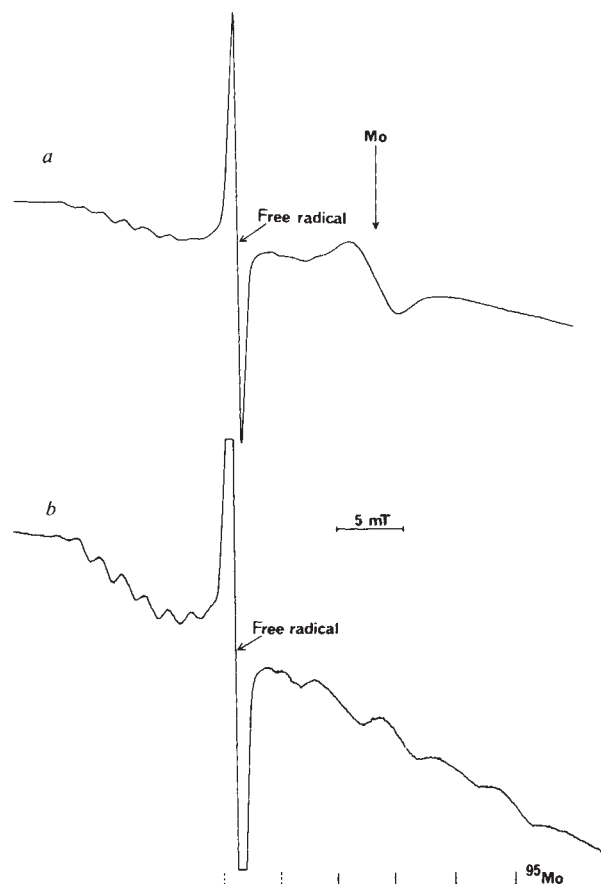


Fig. 2 EPR spectra at ambient temperatures of the Mo(III)-humic acid complex with natural isotopic ratios (*a*) and using ^{95}Mo (*b*).

Mo(I) and Mo(III) would be likely to produce isotropic spectra in the solid state. (Ions with an even number of unpaired electrons, such as Mo(IV), usually require very low temperatures for EPR experiments and even then may not produce a spectrum if the zero-field splitting is large.) While the presence of the former ion would be highly unlikely on chemical grounds, it seems also to be excluded by the EPR parameters because it should have a g -value very close to 2.00, by analogy with other high-spin d^5 ions⁷, and the value of 1.944 (Table 1) obtained from the spectrum is significantly less than this. Thus these EPR results can only be reasonably interpreted as arising from a Mo(III) species.

Note, however, that the intensities of these spectra are very much less (<1%) than those from the initial samples, hence the Mo(III) accounts for only a very small proportion of the molybdenum originally added to the humic acid. Nevertheless, the formation of Mo(III) is highly significant because, although it has been proposed as an intermediate in the reactions of some molybdenum-containing enzymes^{9,10}, there has been little definitive evidence for its formation in natural systems. Also, most complexes of Mo(III) that have been characterized are unstable and readily oxidized by air and mild reducing agents¹¹, but in the present work this ion was formed and remained stable in the solid state even in aerobic conditions. We hope, therefore, that this discovery will provide a stimulus for further searches for the presence of Mo(III) in natural systems.

Received 15 June; accepted 10 August 1982.

1. Beres, T. & Kiraly, I. *Agrokém. Talajt.* **7**, 281–286 (1958).
2. Szilagyi, M. *Soil Sci.* **111**, 233–235 (1971).
3. Goodman, B. A. & Cheshire, M. V. *J. Soil Sci.* **30**, 85–91 (1979).
4. Goodman, B. A. & Cheshire, M. V. *Nature new Biol.* **244**, 158 (1973).
5. Glenworth, R. & Muir, J. W. *Mem. Surv. Scotl.*, Edinburgh (HMSO, 1963).
6. Cheshire, M. V. *et al. J. Soil Sci.* **30**, 315–326 (1979).
7. Goodman, B. A. & Raynor, J. B. *Adv. inorg. Chem. Radiochem.* **13**, 135–362 (1970).
8. Goodman, B. A. & Cheshire, M. V. *J. Soil Sci.* **27**, 337–347 (1976).
9. Ketchum, P. A., Johnson, D., Taylor, R. C., Young, D. C. & Atkinson, A. W. *Adv. Chem. Ser.* **162**, 408–420 (1977).
10. Jacob, G. S. & Orme-Johnson, W. H. in *Molybdenum and Molybdenum-Containing Enzymes* (ed. Coughlan, M.) (Pergamon, New York, 1980).
11. Stiefel, E. I. *Prog. Inorg. Chem.* **22**, 1–223 (1977).

Isotopic variations within a single oceanic island: the Terceira case

Bernard Dupré, Bernard Lambret & Claude J. Allègre

Laboratoire de Géochimie-Cosmochimie (LA 196), Institut de Physique du Globe de Paris et Département des Sciences de la Terre, 4 place Jussieu, 75230 Paris Cedex 05, France

Since the early work of Gast *et al.*¹ it has been recognized that some isotopic variations exist on the scale of a single oceanic island^{2,3}, although the cause of such isotopic heterogeneity remains a matter of debate. For some authors, such as Grant⁴, the variations are related to local contamination in the island. Others, for example, Chase⁵, consider isotopic variation within an oceanic island to reflect mantle heterogeneity, so that it can be used to characterize the mantle and calculate the age of mantle differentiation (mantle isochron). This debate is extremely important in connection with the problem of the isotopic characteristics of the reservoir, the product of which built the oceanic island, and more generally for the structure of the mantle. Terceira is an island of the Azores Archipelago for which we have obtained several kinds of information, including geological⁶, geochronological⁷ and geochemical^{8,9} data, thus enabling us to study its isotopic composition in a well known context. We report here a Sr and Pb isotopic analysis completed on the same samples as were analysed for the major and trace elements¹⁰, and K/Ar dating⁷. Our isotopic results are correlated with the emplacement age of the rocks. This observation is in agreement with a model of mixing between two different mantle components, as has previously been proposed^{2,11,12} for other oceanic islands.

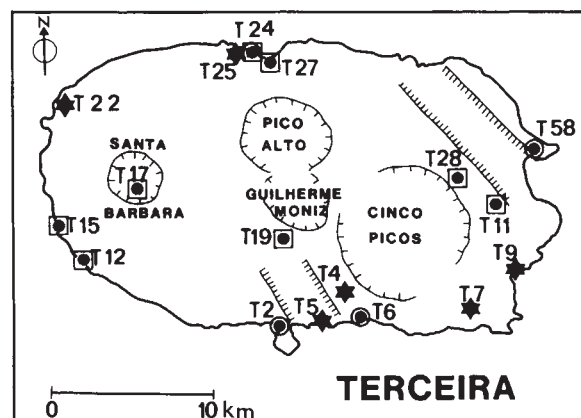


Fig. 1 Localities of samples from Terceira Island. ○, Antecaldera; □, syncaldera; ★, postcaldera.

The volcanic activity on Terceira can be divided into three periods: (1) formation of an antecaldera basaltic shield dated⁷ at 300,000 yr subsequently affected by collapses along a NW–SE orientation; (2) four strato-volcanoes with caldera—Cinco Picos, Santa Barbara, Guilherme Moniz and Pico Alto—were then built; (3) recent volcanic activity produced trachytic domes and basaltic lava flows. The samples analysed are representative of these three units and their location is given in Fig. 1, while Table 1 gives their petrographic characteristics and the Pb and Sr isotopic results. Our results clearly show the existence of important isotopic variations in both Sr and Pb.

The Pb results are plotted on a diagram of $^{206}\text{Pb}/^{204}\text{Pb}$ versus $^{207}\text{Pb}/^{204}\text{Pb}$ (Fig. 3) and show large isotopic variations: $^{206}\text{Pb}/^{204}\text{Pb}$ ratios range from 19.37 to 20.02, $^{207}\text{Pb}/^{204}\text{Pb}$ from 15.57 to 15.64 and $^{208}\text{Pb}/^{204}\text{Pb}$ from 38.95 to 39.35. Sample T27, which is enriched in ^{87}Sr , is not significantly different from the other samples in the Pb diagram.

When examined in a $^{87}\text{Sr}/^{86}\text{Sr}$ versus Sr diagram, the data display isotopic variations of different orders of magnitude: the basalts that are only slightly differentiated exhibit a 4‰ variation (values range from 0.7034 to 0.7038) whereas the highly evolved lavas (pantellerite, trachyte) yield isotopic ratios that are a few ‰ larger than the former (T24 and T27 values are 0.7060 and 0.7075 respectively). Such variation has been observed by Gast *et al.*¹ on Gough and Ascension Islands.

Post-crystallization alteration is often invoked to explain erratic Sr isotopic variations and this has led Dash *et al.*¹³ and O'Nions *et al.*¹⁴ to use leaching techniques to test the influence of this parameter. We have also performed several leaching experiments on the Terceira samples with varying degrees of severity. Our analyses of the residues are listed in Table 2 together with analyses of the leachates which were systematic.

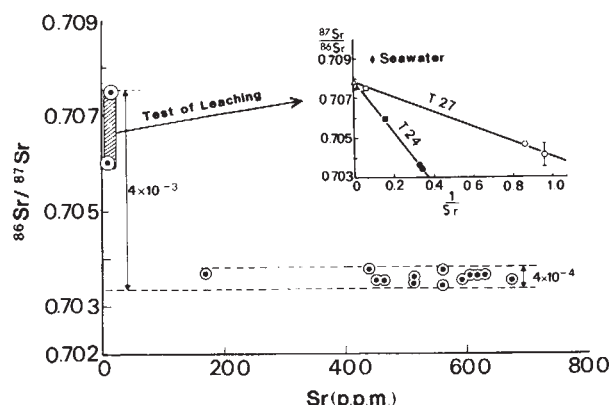


Fig. 2 $^{87}\text{Sr}/^{86}\text{Sr}$ versus Sr diagram: samples with low Sr concentrations yield high Sr-isotopic ratios. Inset: $^{87}\text{Sr}/^{86}\text{Sr}$ versus $1/\text{Sr}$: results of the leaching experiments. T24: ■, whole rocks; ▲, leachates; ●, residues. T27: □, whole rocks, △, leachates; ○, residues.

Table 1 Petrographic characteristics and Pb and Sr isotopic results in samples from Terceira Island

		$^{206}\text{Pb}/^{204}\text{Pb}$	$^{207}\text{Pb}/^{204}\text{Pb}$	$^{208}\text{Pb}/^{204}\text{Pb}$	$^{87}\text{Sr}/^{86}\text{Sr}$	Sr(p.p.m.)	Rb(p.p.m.)	$^{87}\text{Rb}/^{86}\text{Sr}$	
Antecaldera	T2	19.919	15.593	39.07	0.70379 ± 9	440	12.3	0.079	Basalt
					0.70376 ± 11				
	T6	19.930	15.598	39.24	0.70370 ± 12	170	90	1.47	Commendite
	T31				0.70353 ± 7	593	23	0.110	Basalt
	T58				0.70364 ± 7	630	27.7	0.124	Hawaiite
Syncaldera	T77				0.70350 ± 5	452	16.3	0.101	Basalt
	T12	20.022	15.641	39.35					Basalt
	T15	19.861	15.617	39.15					Basalt
	T17	19.973	15.591	39.13	0.70361 ± 6	618	36	0.162	Andesite
	T19	19.799	15.587	39.13					Commendite
	T11	19.969	15.596	39.23	0.70373 ± 6	561	30	0.148	Basalt
	T28	19.748	15.605	39.16	0.70340 ± 9	560	22.3	0.112	Basalt
	T24				0.70600 ± 10	6.3	138	62	Pantellerite
Postcaldera	T27	19.824	15.625	39.29	0.70752 ± 15	17	122	20.1	Trachyte
	T22	19.859	15.575	39.08					Commendite
	T4	19.608	15.570	39.05	0.70347 ± 8	516	24.3	0.132	Basalt
	T25	19.800	15.598	39.15	0.70352 ± 7	676	30.5	0.127	Andesite
	T7	19.726	15.594	39.08	0.70359 ± 7	516	14.7	0.070	Basalt
	T9	19.368	15.570	38.95	0.70352 ± 8	463	21	0.126	Basalt
	T5	19.737	15.575	39.02	0.70362 ± 6	608	35	0.760	Andesite

Pb-Sr compositions have been measured using the technique described by Manhes *et al.*¹⁷ and Birck and Allègre¹⁸ respectively. The Sr results were obtained in 1978 in the course of a study towards a third cycle doctorate. Values for Terceira can be compared with those obtained for samples from the same location by White and Schilling¹⁹. Analytical errors for the $^{206}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$, and $^{208}\text{Pb}/^{204}\text{Pb}$ ratios are 0.02, 0.02 and 0.06 respectively.

cally done when the residue and the whole rock yielded different ratios. These data are presented in Fig. 2 (inset) and display two important points.

First, the high $^{87}\text{Sr}/^{86}\text{Sr}$ ratio values obtained on T24 and T27 may be due to post-crystallization alteration, as the composition of their residue is similar to the composition of the other basalts. The results obtained for the liquid extracted by a weak leaching (ultrasonically processed with cold 6 M HCl) show a Sr isotopic value of 0.708, that is, close to the composition of seawater. Second samples T2, T7 and T17 did not show significant modifications of their isotopic ratios during the leaching experiments, and therefore alteration cannot be invoked in their cases. Furthermore, the existence of a relatively good correlation between Pb and Sr isotopic ratios is another argument against alteration as the cause of the isotopic variations observed among the slightly differentiated basalts (Fig. 3, inset).

During these leaching experiments, we also analysed the lead isotopic composition of the samples and found no significant variation. Post-crystallization alteration thus affects only the end-products of the differentiation in which the Sr concentrations are very low (< 10 p.p.m.). The use of the latter to date the volcanic activity of Terceira seems unjustified.

After discarding alteration, we looked for possible correlations between these variations and other parameters, such as the geographical location, differentiation index or age of emplacement.

(1) No correlation can be clearly established between the geographical location and the composition.

(2) We found no clear correlation on plotting the isotopic results against the differentiation index as defined by Allègre *et al.*⁸, Minister *et al.*⁹ and Provost *et al.*¹¹, using either trace or major elements.

The whole series of samples which seems to define a common trend of fractional crystallization have variable isotopic ratio values and must therefore belong to different magmatic series. It is quite surprising, however, that despite these isotopic differences, they define a common chemical evolution; this demonstrates that when the trace or major element compositions of the starting magma are similar, the differentiation processes yield similar chemical evolution trends which overlap and give the impression of a unique trend. This has also been demonstrated through ^{230}Th - ^{238}U systematics, as defined by Condomines *et al.*¹⁵ for Mt Etna.

In the case of Terceira, an extensive study done by M. Jordan (personal communication) showed that the trace element trend is, in fact, split into several trends; this corroborates our analysis.

Since the Pb or Sr isotopic variations do not seem to be correlated with the chemical differentiation of the volcanic series, it is quite logical that the variations are not produced during the differentiation processes.

(3) A clear relationship exists between the isotopic composition and the age of eruption: the isotopic values of recent basalts are significantly lower than those of older basalts. In addition, the Pb and Sr values of the oceanic crust measured on the ridge at the latitude of the Azores (ref. 16 and B. Hamelin, B. D. and C.J.A. in preparation) are close to those obtained on recent rocks from Terceira (Fig. 3).

Table 2 Sr results obtained by the leaching technique

		Leaching intensity			Whole rocks
		6 M HCl, 20 min ultrasonic 30 °C	6 M HCl, 1 h, ultrasonic 30 °C	6 M HCl, 12 h bombs, 130 °C	
T24	Liquid	$^{87}\text{Sr}/^{86}\text{Sr}$		0.70754 ± 10	
		Sr (p.p.m.)		47	
	Residue	$^{87}\text{Sr}/^{86}\text{Sr}$	0.70358 ± 16	0.70344 ± 9	0.70600 ± 10
		Sr (p.p.m.)	3.04	2.96	
	Liquid	$^{87}\text{Sr}/^{86}\text{Sr}$	0.70785 ± 14	0.70778 ± 9	
		Sr (p.p.m.)	334	165	
T27	Residue	$^{87}\text{Sr}/^{86}\text{Sr}$	0.70467 ± 10	0.70410 ± 6	0.70752 ± 15
		Sr (p.p.m.)	1.16	1.04	
		$^{87}\text{Sr}/^{86}\text{Sr}$		0.70384 ± 7	0.70379 ± 9
					0.70376 ± 11
T2	Residue	Sr (p.p.m.)		159	440
		$^{87}\text{Sr}/^{86}\text{Sr}$		0.70362 ± 6	0.70359 ± 7
	Residue	Sr (p.p.m.)		204	516
		$^{87}\text{Sr}/^{86}\text{Sr}$		0.70357 ± 6	$0.70361 \pm$
T17	Residue	Sr (p.p.m.)			618

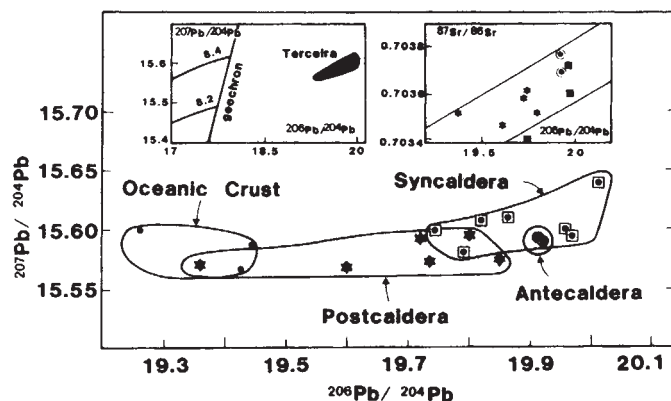


Fig. 3 $^{206}\text{Pb}/^{204}\text{Pb}$ versus $^{207}\text{Pb}/^{204}\text{Pb}$ diagram: the samples are grouped according to their age of eruption. Symbols as in Fig. 1. Upper left corner inset: location of Terceira Island within the (J) field; upper right corner inset: Pb-Sr correlation for Terceira samples.

Such similarities between Pb and Sr isotopic compositions of the oceanic crust and recent alkali basalts and the fact that the isotopic composition changed through time on Terceira indicate that an important exchange occurs between the source of the island basalts and the oceanic crust and/or the source of the oceanic tholeiites. In such a contamination model, the minimum values of the non-contaminated end member-producing Terceira volcanics are 20.05 for $^{206}\text{Pb}/^{204}\text{Pb}$, 15.65 for $^{207}\text{Pb}/^{204}\text{Pb}$ and 39.40 for $^{208}\text{Pb}/^{204}\text{Pb}$ and 0.7038 for $^{87}\text{Sr}/^{86}\text{Sr}$.

As discussed above, the contamination does not seem to occur in the magma chamber, where the chemical differentiation takes place, but rather, in a deeper reservoir. If we admit that blobs (hot spots) are the source material of oceanic islands, it is therefore possible that the contamination occurs during the blob ascension. Similar interpretations have been proposed by Tatsumoto¹² for the Hawaiian Islands and by Sun and Jahn¹¹ for Iceland.

This result is critical because it invalidates the possibility of using the straight lines defined in the $^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$ diagram to infer the age of the differentiation of the mantle source of these basalts as Chase⁵ has done. These straight lines must be considered as mixing lines with no chronometric significance. The Terceira case is by no means unique and our explanation can be extended to St Helena⁴, the Canaries² and certainly to other oceanic islands. As a consequence, only the least contaminated member may be considered representative of the hot spot source. For Terceira this is most probably on the (J) field in the Pb-Pb diagram, in agreement with the general characteristics of oceanic islands (Fig. 3, inset).

We thank C. Gariépy and C. Mercier for help. This is IPGP contribution NS/605.

Source of the detrital components of uraniferous conglomerates, Quirke ore zone, Elliot Lake, Ontario, Canada

Andrew Robinson* & Edward T. C. Spooner

Department of Geology, University of Toronto, Toronto, Ontario, Canada M5S 1A1.

Large volumes of pyritic quartz-pebble conglomerate in the Elliot Lake district of Ontario contain at least 300,000 tonnes U_3O_8 recoverable at a grade of 0.1% (ref. 1), with much greater tonnages available at lower recovery grades. The Quirke zone alone contains in excess of 224,000 tonnes U_3O_8 at ~0.1% (refs 2-4) and is, accordingly, one of the world's largest known uranium deposits. Situated 40 km north of the North Channel of Lake Huron, 400 km north-west of Toronto, ore conglomerates occur in the Matinenda Formation, locally the basal unit of the Huronian Supergroup, and, although their age is not well constrained, they are probably close to 2,350 Myr old.⁵ Uraninite, the most important primary ore mineral, occurs as poorly rounded to euhedral grains 0.05-0.2 mm across within the matrix between quartz pebbles. While its chemical and morphological characteristics have led most workers to consider that the uraninite is detrital⁵⁻⁷, there has hitherto been no attempt to characterize its source lithology, a matter of considerable significance as regards mineral exploration. This was the aim of the present study. We have examined the mineral and fluid inclusions observed in the quartz pebbles in the conglomerate and conclude that they were derived from a pegmatitic, potassic (alaskitic) granite. We have also demonstrated that in this source lithology, pebble-sized quartz was spatially associated with radioactive minerals, by inference, uraninite. The postulated characteristics of such a source are remarkably similar in several respects to the uraniferous alaskites of the Rössing deposit in Namibia.

At least three features of clastic sediments can provide clues concerning the nature of their source lithologies: (1) internal textures of detrital grains; (2) sediment bulk mineralogy and chemistry; (3) mineral and fluid inclusions in detrital grains. In each case, it is necessary to distinguish features possibly inherited from source lithologies from those that developed subsequent to deposition and within the sediment. The Elliot Lake conglomerates have been significantly deformed through pressure solution⁷ and although many of the deformation textures displayed by quartz clasts can be demonstrated to have originated after deposition, we have been unable to identify such textures that might have been inherited from a source lithology, as sometimes is possible in undeformed sediments⁸. Table 1 lists the minerals present in the conglomerate along with their presumed origins (that is, detrital or authigenic).^{5,7} The detrital mineral suite suggests, in a general way, a source of alaskitic affinities. However, because clastic components of sediments are frequently derived from more than one precursor, the best method of characterizing the source of any particular mineral is to examine its mineral and fluid inclusions. The uraninite grains themselves occasionally contain inclusions of quartz but these are invariably associated with secondary coffinite and are considered to be mostly intrastratal replacement features. Even if some quartz inclusions are inherited, these cannot be distinguished. We have therefore examined mineral and fluid inclusions in ore horizon quartz.

Mineral inclusions that have been introduced into quartz clasts during pressure solution deformation in the conglomerate are readily identifiable through their concentration around pebble margins or by the 'trail' of sericite left in their lee (Fig.

Received 1 June; accepted 26 July 1982.

- Gast, P. W., Tilton, C. R. & Hedge, C. E. *Science* **145**, 1181 (1964).
- Sun, C. S. *Phil. Trans. R. Soc. A297*, 409 (1980).
- Hofmann, A. W. & Hart, S. R. *Earth planet. Sci. Lett.* **38**, 44 (1978).
- Grant, N. K., Powell, J. L., Burkholder, F. R., Walther, J. V. & Coleman, M. L. *Earth planet. Sci. Lett.* **31**, 209 (1976).
- Chase, G. G. *Earth planet. Sci. Lett.* **52**, 277 (1981).
- Zbyszewsky, G. *Mem. Acad. Cien. Lisboa* **12**, 185 (1968).
- Feraud, G., Kaneoka, I. & Allègre, C. J. *Earth planet. Sci. Lett.* **46**, 275 (1980).
- Allègre, C. J., Treuil, M., Minster, J. F., Minster, B. & Albarède, F. *Contr. Miner. Petrol.* **60**, 57 (1977).
- Minster, J. F., Minster, J. B., Treuil, M. & Allègre, C. J. *Contr. Miner. Petrol.* **61**, 49 (1977).
- Provost, A. & Allègre, C. J. *Geochim. cosmochim. Acta* **43**, 487 (1979).
- Sun, S. S. & Jahn, B. M. *Nature* **255**, 527 (1975).
- Tatsumoto, M. *Earth planet. Sci. Lett.* **38**, 63 (1978).
- Dasch, E. J., Hedge, C. E. & Dymond, J. *Earth planet. Sci. Lett.* **19**, 177 (1973).
- O'Nions, R. K. & Pankhurst, R. J. *Earth planet. Sci. Lett.* **31**, 255 (1976).
- Condomines, M. *Geochim. cosmochim. Acta* (in the press).
- Dupré, B. & Allègre, C. J. *Nature* **286**, 17 (1980).
- Manhes, G., Minster, J. F. & Allègre, C. J. *Earth planet. Sci. Lett.* **39**, 269 (1978).
- Birck, J. L. & Allègre, C. J. *Earth planet. Sci. Lett.* **39**, 37 (1978).
- White, N. M., Schilling, J. G. & Hart, S. R. *Nature* **263**, 659 (1976).

* Present address: Billiton International Metals, BV, 19 Louis Couperusplein, Box 190, The Hague, The Netherlands.

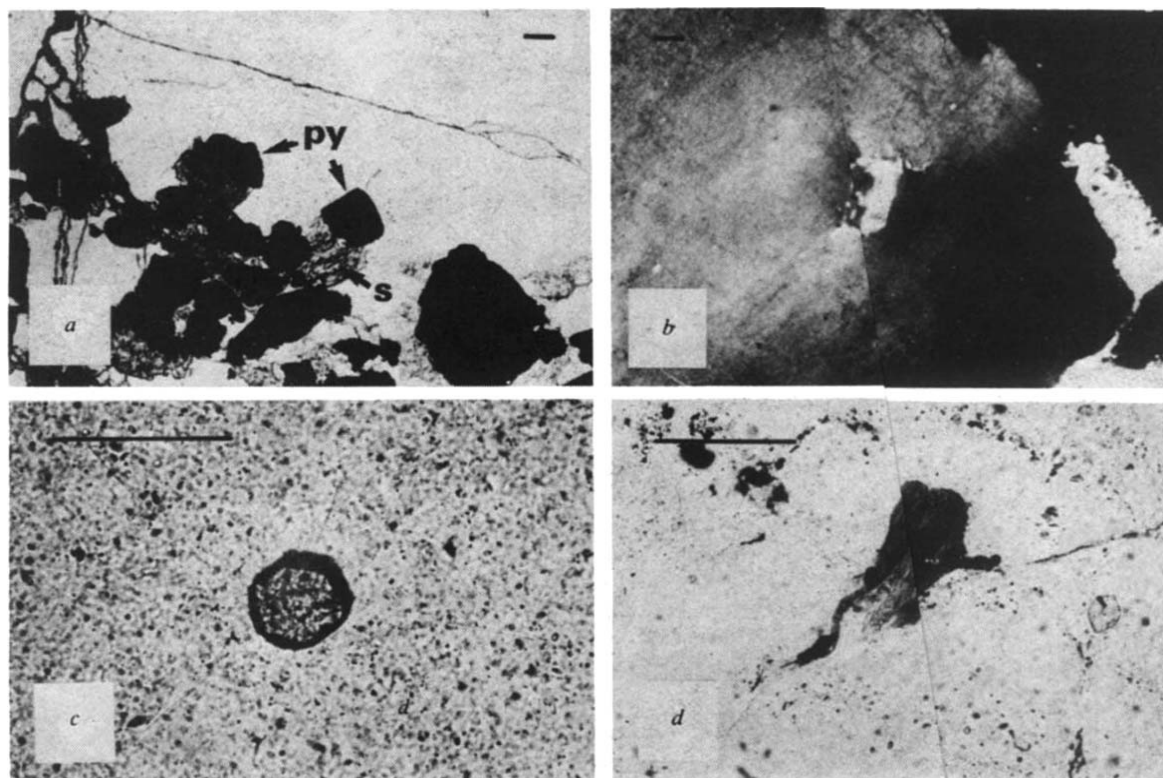


Fig. 1 Mineral inclusions in quartz clasts. *a*, Pyrite euhedra (py) introduced during pressure-solution deformation within the conglomerate. A trail of sericite (s) commonly occurs behind such inclusions. *b*, Euhedral, primary orthoclase inclusion. *c*, Euhedral, primary zircon inclusion. *d*, Altered biotite inclusion. Scale bar, 0.1 mm.

1a). Primary (inherited) inclusions larger than 1–2 μm across are rare. We have identified microcline—the most common, often intergrown with quartz as lithoclasts—orthoclase, zircon and biotite (Fig. 1b–d). In addition, minute yellow needles, abundant in certain clasts, are probably rutile. Although plagioclase has not been recognized as an inclusion, it is very rare as a detrital phase. Neither have uraninite crystals been found in quartz clasts. This mineral suite is very characteristic of alkali-feldspar granites (alaskites) and the grain size of the quartz pebbles (generally <6 cm, but up to 20 cm across) indicates that such a granite would have been in part pegmatitic. In particular, zircon is not known to occur as a hydrothermal mineral in veins, such as have previously been advanced as possible sources for the quartz⁵.

Drawing a distinction between primary and post-depositional features is particularly difficult in the study of fluid inclusions

in quartz and we have therefore subdivided these non-genetically (Table 2). Some type Ia and pe IIa inclusions are demonstrably post-depositional in origin: the healed fractures along which they occur can sometimes be traced along their length into other clasts or into sericite-filled fractures. Similarly, the type Ib and IIb inclusions that occur in the rare quartz cement must have been entrapped following sedimentation. The fluid inclusions that are most likely to have been inherited from the source from which the quartz pebbles were eroded are those that contain liquid CO_2 (type IV) or those relatively large (up to 8 μm across) clustered, highly saline inclusions that contain a variety of, and usually several, included mineral phases (type IIb). Although both these types of inclusion are generally rather rare, a suite of this kind is characteristic of granitic rocks in general, including certain pegmatitic alaskites^{9,10}, and is not typical of hydrothermal veins.

The discussion has concentrated on characterizing the lithology that supplied quartz to the conglomerates. This is only relevant to a study of the source of uraninite if it can be demonstrated that the two minerals are co-genetic. A characteristic of ore conglomerates is the common occurrence of smoky quartz pebbles, the smokiness being a feature generally attributed to the formation of free silica through exposure to radioactive minerals¹¹. Such pebbles are assumed by most workers to have been irradiated after deposition and while this may well have been the case in some instances, several ore samples have been found in which distinctly smoky and milky, or clear, quartz pebbles are juxtaposed (Fig. 2). Serial sections were cut every 8 mm across two such samples from Elliot Lake (Fig. 2a,b) and autoradiographs of the sections revealed, in both cases, no preferential association of radioactivity with the smoky pebble. For comparison, the same experiment was performed on a single similar sample from the Steyn Placer of the Upper Witwatersrand System of South Africa with the same result (Fig. 2c). Smoky quartz pebbles may also occasionally be observed in the very poorly radioactive 'IQ' (intermediate quartzite) between ore reefs. The conclusion that follows is that smokiness

Table 1 The mineralogy of the Elliot Lake conglomerates

Mineral	Abundance	Origin
Quartz	A	D (+ minor cement)
Microcline	A	D
Sericite	A	Au (after feldspar) + some D(?)
Pyrite	A	Au + minor D
Uraninite	C	D
Rutile-brannerite	C	Au (after detrital magnetite-ilmenite)
Monazite	C	D
Zircon	C	D
Uranothorite	R	Au + D
Coffinite	R	Au
Yttrium-uranium-REE-phosphate	R	Au
Chromite	VR	D
+ many other very rare detrital heavy minerals		

A, abundant; C, common; R, rare; VR, very rare; D, detrital; Au, authigenic. Sources: refs 5–7.

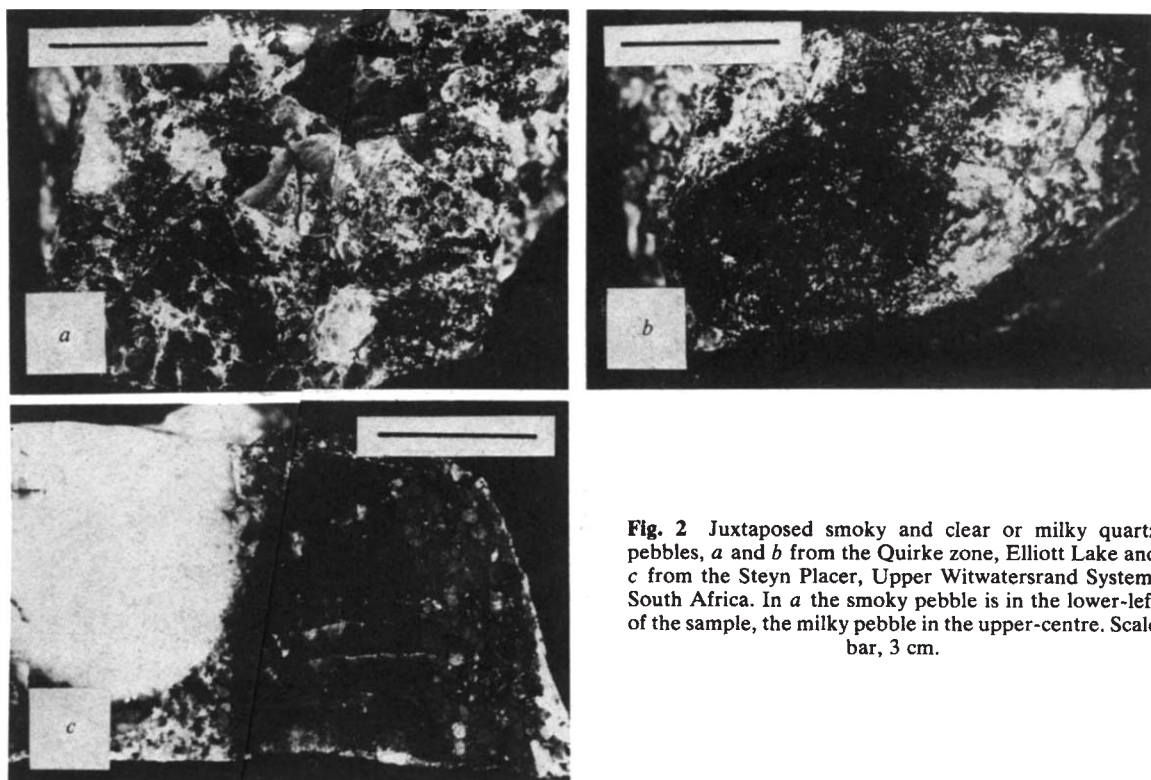


Fig. 2 Juxtaposed smoky and clear or milky quartz pebbles, *a* and *b* from the Quirke zone, Elliott Lake and *c* from the Steyn Placer, Upper Witwatersrand System, South Africa. In *a* the smoky pebble is in the lower-left of the sample, the milky pebble in the upper-centre. Scale bar, 3 cm.

is at least in part a feature inherited from a source lithology in which coarse quartz was associated with radioactive minerals. The most important radioactive detrital mineral preserved in the ore conglomerates is uraninite. (Although monazite may contain substantial amounts of thorium and uranium, these are mainly alteration products that originated during the diagenesis of the conglomerates⁷. Monazite may nonetheless have been fairly radioactive within the source lithology.)

The evidence described above suggests that the major detrital components of the Elliott Lake uraniferous conglomerates were derived from a uraninite-bearing, pegmatitic, alaskitic granite. A comparison can be drawn between this postulated source lithology and the Rössing deposit in Namibia, where Lower Palaeozoic pegmatitic askites are being mined for uranium^{12,13}. The pegmatites contain major smoky quartz and perthitic microcline with an overall typical grain size of 0.5–5 cm. There is minor orthoclase, plagioclase and biotite and a

suite of accessory minerals, the most common of which are zircon and monazite. Uraninite, the main primary ore mineral, concentrates with the potassic minerals microcline and biotite. Both its range of grain sizes (0.05–0.1 mm) and its ThO₂ content (~7%) are remarkably similar to the corresponding characteristics of Elliott Lake uraninite grains (grain size 0.05–0.2 mm, ThO₂ content 4–8%)⁶. Clustered, highly saline, as well as CO₂-rich fluid inclusions are characteristic of the smoky quartz in the pegmatites¹⁰. The Rössing deposit contains ~136,000 tonnes U₃O₈ recoverable at a grade of 0.035% (ref. 14). If a deposit of this type and size was eroded and quantitatively deposited as a sediment, a placer uranium deposit with tonnage, grade and petrological characteristics comparable with those of the Quirke zone, Elliott Lake, would have resulted.

A.R. gratefully acknowledges the receipt of a University of Toronto Connaught Scholarship, a Rhodes Scholarship to Canada and an Ontario Graduate Fellowship. E.T.C.S. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada, operating grant A6114. We also thank the staff of the Rio Algom Quirke Mine and the Denison Mine, particularly Russ Gunning; Brian O'Donovan (photographer); and fax (typing services).

Received 13 April; accepted 27 July 1982.

- Boyle, R. W. *Geol. Surv. Can. Bull.* **280**, (1979).
- McMillan, R. H. in *Geology, Mining and Extractive Processing of Uranium* (ed. Jones, M. J.) 43–55 (Institute of Mining and Metallurgy, London, 1976).
- Canadian Mines Handbooks 1977–78 to 1981–82* (Northern Miner Press, Toronto, 1977–81).
- Rio Algom Limited Annual Report, 1980* (1981).
- Roscoe, S. M. *Geol. Surv. Can. Pap.* 68–40 (1969).
- Theis, N. J. *Geol. Surv. Can. Bull.* **304** (1979).
- Robinson, A. G. thesis, Univ. Toronto (1982).
- Young, S. W. *J. Sedim. Petrol.* **46**, 595–603 (1976).
- Konnerup-Madsen, J. *Lithos* **12**, 13–23 (1979).
- Cuney, M. *Trans. geol. Soc. S. Afr.* **83**, 39–45 (1980).
- Hurlbut, C. S. & Klein, C. *Manual of Mineralogy*, 19th edn (Wiley, New York, 1977).
- Berning, J., Cooke, R., Hiemstra, S. A. & Hoffman, U. *Econ. Geol.* **71**, 351–368 (1976).
- Von Backström, J. W. & Jacob, R. E. *Phil. Trans. R. Soc. A291*, 307–319 (1979).
- Nash, J. T., Granger, H. C. & Adams, S. S. *Econ. Geol. 75th Anniversary Vol.* 63–116 (1981).

Table 2 Fluid inclusions quartz in the Elliott Lake conglomerates

Type	Characteristics	Comments
I	H ₂ O ^L + small V bubb	Up to 5 µm across; Ia occurs in both detrital and authigenic quartz; Ib most common type in detrital quartz
II	H ₂ O ^L + small V bubb + contained mineral	IIa contain a small, usually prismatic daughter mineral and may be genetically with most Ia inclusions. Most IIb inclusions are very similar to Ib. Rather rare, clustered, highly saline inclusions with a variety of included solid minerals, up to 8 µm across, are probably not genetically related to type Ib
III	H ₂ O ^L + large V bubble	IIIa and IIIb inclusions are very small (2 µm across) and thus very difficult to characterize
IV	CO ₂ ^L ± H ₂ O ^L + Vubble ± contained mineral(s)	Up to 8 µm across, variable characteristics. Rather rare
V	Highly irregular	Up to 10 µm long; may represent decrepitated inclusions.

Suffix a after the type number indicates that inclusions are fracture related; suffix b indicates that they are isolated.

Discordant layering relations in the Fongen-Hyllingen basic intrusion

J. R. Wilson & S. Brink Larsen

Department of Geology, University of Aarhus,
8000 Århus C, Denmark

Modal stratification or layering on the scale of centimetres to metres is a common macroscopic feature of many basic intrusions. In some cases this rhythmic layering is accompanied by gradual changes in the composition of mineral solid solutions (cryptic layering) and by the entry and exit of certain minerals (phase layering). These compositional variations, which evidently reflect progressive fractional crystallization, are generally assumed to be orthogonal to rhythmic layering, as predicted by crystal settling models for the origin of layered basic intrusions, and they have been used to establish 'cumulate stratigraphy' of many layered intrusions¹. In the southern part of the synorogenic Fongen-Hyllingen complex², Norway, well developed rhythmic layering is markedly discordant to phase and cryptic layering. Regressions in the cryptic and phase layering that are believed to reflect magma addition are also discordant to the rhythmic layering. These features raise questions as to the stratigraphical significance of phase layering and add to the mounting evidence in favour of *in situ* crystallization for the origin of layered igneous rocks^{3,4}.

The 160 km² Fongen-Hyllingen basic complex² (Fig. 1) is a strongly differentiated layered intrusion situated in the Trondheim region of Norway (11.5° E, 63° N). In spite of being emplaced synorogenically before the culmination of the Caledonian regional deformation and metamorphism⁵, it exhibits sufficient primary features to allow detailed study of the igneous mineralogy and petrology. The present form of the

complex is largely controlled by its proximity to the Tydal synform, a late Caledonian fold⁶ (Fig. 1b).

The intrusive complex consists predominantly of rhythmically layered rocks, sections with thicknesses of ~6,200 and 3,600 m being preserved in the northern (Fongen) and southern (Hyllingen) areas, respectively. The stratigraphically lowest rocks (dunites and troctolites with olivine Fo₈₆ and plagioclase An₈₀) occur in the extreme north. The highest, most evolved rocks (quartz-bearing ferrosyenites with Mg-free mafic minerals and albite) occur in the south along the eastern margin of the Hyllingen area. A broadly tholeiitic parent magma apparently fractionated under moderate, increasing P_{H_2O} , the fractionation trend being broadly intermediate between tholeiitic and calc-alkaline trends^{2,7}. Regressions in cryptic variations and phase layering with stratigraphical height suggest periodic magma addition into a chamber that was probably closed to the removal or exit of magma².

The strike of centimetre-metre scale rhythmic layering in most of the structurally simple Hyllingen area⁸ swings gently from ~N15° E in the north to N25° W in the south, broadly parallel with the lower, western boundary (Fig. 1c). Younging is consistently eastwards and dips average ~40° E. A narrow zone of unlayered, broadly gabbroic rocks occurs along the western margin, whereas in the east, syenitic differentiates vein country rock amphibolites with no intervening marginal facies. The rhythmic layering strikes directly towards the southern margin and there is no evidence of local banking of the layering against the wall of the intrusion. Approaching the southern margin, the intensity of rhythmic layering decreases over a distance of ~150 m until, in the immediate vicinity of the contact, layering is absent. The presence of many country rock hornfels inclusions (pendants?) and zones of ferrodioritic and ferrosyenitic differentiates entirely surrounded by hornfels in the extreme south-east confirm that the original wall and/or roof of the intrusion are preserved here.

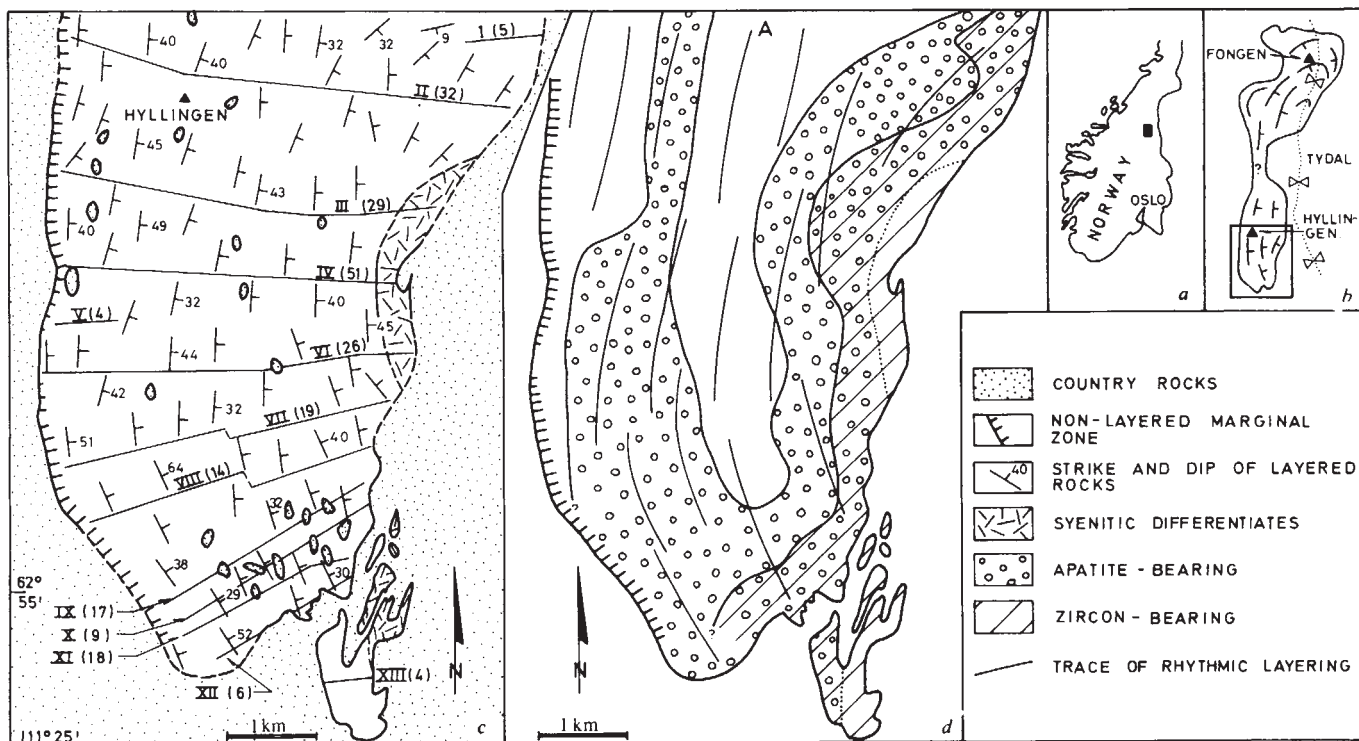


Fig. 1 a, Map showing location of b in Norway. b, Outline map of the Fongen-Hyllingen complex after Wilson *et al.*² showing generalized orientation of the rhythmic layering and the trace of the hinge zone of the Tydal synform. Location of c and d is boxed. c, The southern part of the Fongen-Hyllingen complex. Rhythmic layering strikes directly towards the southern intrusive contact. A thickness of about 2,700 m is exposed in profile IV. Number of samples from each profile is shown in parentheses (total of 234 samples). Average sample density is about 1 per 110 m except for profile IV (1 per 53 m). Profile IV is essentially equivalent to profile IIIC in Fig. 3 of ref. 2. Faulting is minor and has no significant influence on the feature described here. d, Distribution of apatite and zircon in the Hyllingen area based on samples from profiles I–XIII in c. Note strong discordance between the curves showing the entry/exit of these phases and the trace of the rhythmic layering. Detailed sampling has shown that apatite is absent in the final differentiates, in the uppermost ~100 m of profile IV. The final, most fractionated liquid must have been almost completely depleted in phosphorus.

In a profile normal to the strike of the rhythmic layering ending in the most fractionated rocks along the eastern margin, such as profile IV in Fig. 1c (which has a stratigraphical thickness of ~2,700 m) the early crystallizing phases in the lowest layered rocks are olivine (Fo₄₀), plagioclase (An₄₅), Ca-poor pyroxene, Ca-rich pyroxene and Fe-Ti oxides. The sequence of entry (+) and exit (-) of early crystallizing phases in this profile is apatite (+), apatite (-), calcic amphibole (+), apatite (+), biotite (+), zircon (+), quartz (+), Ca-poor pyroxene (-), K-feldspar (+), olivine (-), allanite (+) and apatite (-). The areal distribution of certain phases has been established through petrographical study of 234 samples from 13 profiles (Fig. 1c, d). Lines defining the entry and exit of cumulus apatite and cumulus zircon are markedly discordant to the rhythmic layering and appear to be more closely related to the exposed shape of the intrusion. For example, cumulus zircon enters at a stratigraphical level some 1,200 m lower in profile XI than in profile VI, these profiles being from ~2,300 to 3,000 m apart. Following the strike of the rhythmic layering from point A in Fig. 1d southwards, the compositions of plagioclase and olivine change from about An₆₀ and Fo₇₃ in profile II to about An₃₀ and Fo₁₀ in profile XII. Thus the rocks become increasingly differentiated along the strike of the rhythmic layering as the margin is approached. It would appear that differentiation varied from place to place within the intrusion and was probably influenced by the distance from the roof and wall. Consequently concordance between cumulate stratigraphical zones and rhythmic layering, implicit in the use of phase layering as stratigraphically significant, akin to zone fossils in sedimentary sequences, cannot be assumed for layered intrusions. Similar features have been indicated for the Skaergaard intrusion^{3,9} and favour a model involving at least partly diffusion-controlled *in situ* crystallization rather than crystal settling for the origin of the Fongen-Hyllingen complex.

Abrupt regressions in phase and/or cryptic layering are commonly ascribed to the influx of fresh magma. Such a reversal is represented, for example, in the lower part of profile III by the exit of apatite accompanied by composition changes in plagioclase from about An₄₀ to An₅₅ and in olivine from about Fo₂₀ to Fo₅₀ (ref. 2 and unpublished results). However, work in progress has revealed that this major regression is not an abrupt event. After the disappearance of apatite in profile III, the compositions of solid solution minerals gradually become less evolved over a thickness of some 300 m, after which they again begin to evolve to lower temperature compositions. The regression indicated by the exit of apatite appears more or less concordant with the rhythmic layering in the north but becomes increasingly discordant to the south (Fig. 1d).

Discordance between rhythmic layering and the phase and cryptic layering raises the question as to which, if any, of them represents a constant time surface during crystallization. Compared with the time it takes for a large layered intrusion to crystallize, magma replenishment is generally believed to be a relatively rapid process¹⁰. If this is so, then discordant relations between a major reversal and rhythmic layering imply that the plane of rhythmic layering probably does not represent a constant time surface. Alternatively, the nature of this discordance in the southern Hyllingen area (Fig. 1d) could indicate that replenishment was a slow process, the magma chamber being gradually filled from a direction with a northerly component.

The Carlsberg Foundation of Denmark provided financial support for field work. Dr W. J. Wadsworth criticized the manuscript.

Received 26 April; accepted 23 July 1982.

1. Wager, L. R. & Brown, G. M. *Layered Igneous Rocks* (Oliver & Boyd, Edinburgh, 1968).
2. Wilson, J. R., Esbensen, K. H. & Thy, P. *J. Petrol.* **22**, 584-627 (1981).
3. McBirney, A. R. & Noyes, R. M. *J. Petrol.* **20**, 487-554 (1979).
4. Campbell, I. H. *Lithos* **11**, 311-323 (1978).
5. Olesen, N. Ø., Hansen, E. S., Kristensen, L. H. & Thyrsted, T. *Leid. geol. Meded.* **49**, 259-276 (1973).
6. Wilson, J. R. & Olesen, N. Ø. *Norsk. geol. Tidsskr.* **55**, 423-439 (1975).
7. Wilson, J. R., Esbensen, K. H. & Thy, P. *Nature* **290**, 325-326 (1981).
8. Nilsen, O. *Norsk. geol. Tidsskr.* **53**, 213-231 (1973).
9. Naslund, H. R. Yb. *Carnegie Instn. Wash.* **75**, 640-644 (1976).
10. Irvine, T. N. in *Physics of Magmatic Processes* 325-383 (Princeton University Press, 1980).

Arsenic in Napoleon's wallpaper

David E. H. Jones

Department of Physical Chemistry, The University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK

Kenneth W. D. Ledingham

Department of Natural Philosophy, University of Glasgow, Kelvin Laboratory, East Kilbride, Lanarkshire G75 0QU, UK

Napoleon's illness and death on St Helena have been the subject of controversy. His symptoms have been compared with those of arsenic poisoning¹⁻³, and arsenic has been found in surviving samples of his hair⁴⁻⁸. During the nineteenth century, many people were accidentally poisoned by arsenical vapours from wallpaper⁹⁻¹¹. Accordingly, it is an intriguing speculation that Napoleon may have ingested arsenic from his wallpaper. After presenting this speculation on a radio broadcast, one of us (D.E.H.J.) was offered an original sample of wallpaper from Napoleon's residence on St Helena. We report here that X-ray fluorescence measurements on this paper reveal enough arsenic to be capable of causing illness but probably not death. So if Napoleon was poisoned by arsenic, it may have come not from deliberate administration but from his wallpaper.

After Napoleon Bonaparte's death in May 1821, autopsy revealed an enlarged liver and a stomach lesion¹². However, later neutron activation analyses of his hair⁴⁻⁸ have found abnormal quantities of arsenic in some (but not all) samples from his period on St Helena. Forshufvud and Weider have argued that he was the victim of a poisoning conspiracy¹⁻³, but arsenic was widely used for many medical, cosmetic and environmental applications during the nineteenth century, so he may have ingested it from some quite innocent source.

One possible source is paint or wallpaper. The green copper arsenite pigments (Scheele's green, and Paris or emerald green) were introduced in about 1780, and were soon widely used in paints and wallpapers. By 1815 they were known to be poisonous, and throughout the nineteenth century many people were made ill or even killed by arsenic from their wallpaper⁹⁻¹¹. Not until 1983 was the mechanism of the process elucidated¹³. On damp wallpaper, many moulds (for example, *S. brevicaulis*) can metabolize arsenic compounds to volatile, poisonous arsenic trimethyl, which is released into the air of the room.

Our sample of Napoleon's wallpaper had been preserved in an old family scrapbook. It measured about 55 × 60 mm, with a beige coating carrying a 40-mm diameter flock-decorated rosette, mainly brownish but with bright green areas (Fig. 1).

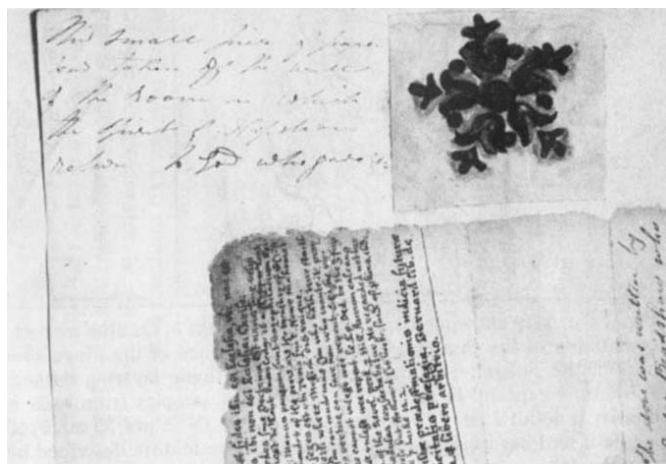


Fig. 1 A sample of Napoleon's wallpaper preserved in an old family scrapbook.

Table 1 Metal content of historical wallpapers, determined by X-ray fluorescence spectroscopy

Element	X-ray observed	Metal content (as the element) (g m^{-2})				
		Napoleon's wallpaper			Craggside papers	
		a	b	c	1	2
Cu	K α	4.6	0.03	0.28	3.8	3.5
As	K α	1.5	0.04	0.12	6.5	6.2
Pb	L β_1	19	0.3	1.3	1.9	0.4
Ca	K α , K β	6.6	14	14	7.5	17
Fe	K α	0.15	0.4	0.4	0.07	0.06

a, Rosette area; b, beige area surrounding rosette; c, mean composition of the original wallpaper, deduced by weighting the compositions of rosettes and beige background according to their relative areas.

It had been removed from the drawing room of Longwood House (the room in which Napoleon died) and pasted into the scrapbook in May or June of 1825. At the time, the room had a cream wallpaper with blue stars¹⁴; the brownish rosette must be one of the stars, faded over the years. For comparison, other samples of nineteenth century wallpapers were sought from the National Trust. Two specimens of 1864, from 'Craggside', Lord Armstrong's country house in Northumberland, proved highly arsenical.

The wallpapers were analysed primarily by X-ray fluorescence spectroscopy (XRF). Our equipment used a ^{147}Pm source of 1.5×10^9 Bq, a Si(Li) detector of area 80 mm^2 , and a 2,000-channel analyser covering the range 0–60 keV; all elements from potassium to lead show either K or L X-ray fluorescence within this range. The spectra of the wallpapers were dominated by peaks due to copper, arsenic, lead, calcium and iron.

The spectra were calibrated by comparing them with those of calibration papers made up to contain known amounts of these metals. Interpretation was complicated by the presence of lead pigments in the papers. The L α line of lead almost coincides with the K α line of arsenic, and the lead components had to be subtracted from the recorded spectra before the arsenic peaks could be measured. Furthermore, a dense layer of lead pigment in the rosette of Napoleon's wallpaper attenuated the radiation from the metal pigments beneath it. They were estimated by reversing the paper in its mount and irradiating it from the back. The final results are shown in Table 1.

Matrix and attenuation effects, inevitable in the XRF analysis of heterogeneous samples, make the results uncertain by ± 20 –30%. Within these limits, the XRF results are supported by neutron and photon activation analyses of pigment flakes detached from Napoleon's wallpaper, and chemical analysis of a section of Craggside 1. The average composition of Napoleon's wallpaper (Table 1) was derived by taking the rosettes to be 12 cm^2 in area, and spaced¹⁴ at one per 225 cm^2 on the beige background. On this basis, the rosettes contribute 0.08 g m^{-2} of arsenic, and the beige background 0.04 g m^{-2} ; a total of 0.12 g m^{-2} .

With this fairly mild arsenic loading, Napoleon's wallpaper could not have posed a serious threat to his life, but it may well have contributed to his illness. In suitably damp and mouldy conditions, highly arsenical wallpapers like Craggside 1 and 2 could cause illness or even death in a few days or weeks^{15–17}, but even much less arsenical papers could cause notable ill-health. In 1893, a detailed study¹¹ showed that wallpapers containing 0.6 – 0.015 g m^{-2} of arsenic could cause illness, and suggested that even 0.006 g m^{-2} could be hazardous. The safety standards of the time¹⁸ recommended 0.001 – 0.005 g m^{-2} as a safe upper limit. Thus, our finding of 0.12 g m^{-2} of arsenic in Napoleon's wallpaper is toxicologically significant, particularly in view of the arsenic found in his hair by several neutron-activation studies.

These studies, on hair samples spanning 1816–21, reported irregularly fluctuating levels of arsenic, as would be expected from a source like damp and mouldy wallpaper. Fluctuations with humidity, temperature, ventilation, and the amount of

time that Napoleon spent in the relevant rooms, would be expected. Longwood was notoriously damp¹⁹, and many of its other inhabitants also suffered from its 'unhealthy atmosphere'¹.

The present sample of Napoleon's wallpaper was only installed¹⁴ in 1819, so cannot have contributed to his arsenic intake before that time. This report does, however, demonstrate that hazardously arsenical wallpaper was available on St Helena during Napoleon's captivity, and in one instance at least was used to decorate his residence. Accordingly, conspiracy theories need not be invoked to explain arsenic found in his hair.

Received 17 May; accepted 10 August 1982.

- Forshufvud, S. *Who Killed Napoleon?* (Hutchinson, London, 1962).
- Forshufvud, S. & Weider, B. *Assassination at St Helena* (Mitchell, Vancouver, 1978).
- Weider, B. & Hapgood, D. *The Murder of Napoleon* (Robson Books, London, 1982).
- Smith, H., Forshufvud, S. & Wassen, A. *Nature* **194**, 725–726 (1962).
- Forshufvud, S., Smith, H. & Wassen, A. *Archs Tox.* **20**, 210–219 (1964).
- Leslie, A.C.D. & Smith, H. *Archs Tox.* **41**, 163–167 (1978).
- Forshufvud, S., Smith, H. & Wassen, A. *Nature* **192**, 103–105 (1961).
- Lewin, P. K., Hancock, R. G. V. & Voynovich, P. *Nature* **299**, 627–628 (1982).
- Challenger, F. *Chem. Rev.* **36**, 315–361 (1945).
- Hunter, D. *Diseases of Occupation*, 348, 363 (Hodder & Stoughton, London, 1978).
- Sanger, C. R. *Proc. Am. Acad. Arts Sci.* **29**, 148–177 (1893).
- Hillemand, P. *Pathologie de Napoleon*, 119–187 (La Palatine, Paris, 1970).
- Gosio, E. *Archs ital. Biol.* **18**, 235–298 (1893); *Ber. dt. chem. Ges.* **30**, 1024–1026 (1897).
- Longwood Old House* (official pamphlet) 8–9 (1960).
- Huss, H. Z. *Hyg.* **76**, 361 (1914).
- Metcalfe, J. B. *Lancet* **ii**, 535–536 (1860).
- Orton, T. *Lancet*, **ii**, 516–518 (1862).
- Brunton, T. L. B. *med. J. I*, 1218–1221 (1883).
- Korngold, R. *The Last Years of Napoleon*, 84–85 (Harcourt, Brace & Co., New York, 1959).

Napoleon Bonaparte—no evidence of chronic arsenic poisoning

Peter K. Lewin, Ronald G. V. Hancock* & Paul Voynovich

The Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8

* Slowpoke Reactor Facility, University of Toronto, Toronto, Ontario, Canada M5S 1A4

Napoleon Bonaparte died in exile on the island of St Helena in 1821 at the age of 51. An autopsy by his physician, Autommarchi, in the presence of Sir Thomas Reade, some staff officers and eight medical men revealed a cancerous growth of his stomach¹. But mainly on the basis of high levels of arsenic detected in several samples of his hair² it has been suggested, most recently by Hapgood and Weider³, that Napoleon was poisoned by arsenic. We have analysed a different sample of Napoleon's hair and find an almost normal arsenic content but elevated antimony.

A lock of hair was obtained by Major Poppleton, one of the staff officers on St Helena. Major Poppleton gave this hair sample to his uncle, whose relatives retained it until the First World War (Fig. 1). It was then obtained by the late Mr Gilbert William Plenty, who gave it to his son, Mr Albert William Plenty of Toronto, in whose custody it now is. The specimen consists of a well preserved lock of hair, chestnut in colour (Fig. 1). Gross and microscopic examination have revealed no structural abnormalities.

A sample (12 mg) of the hair was placed in a 1 ml polyethylene vial and irradiated in the Slowpoke Nuclear Reactor of the University of Toronto for 16 h at a thermal neutron flux of $2.5 \times 10^{11} \text{ N cm}^{-2} \text{ s}^{-1}$ for neutron activation analysis.

After a 7-h delay to allow short-lived radioisotopes to decay away, the hair was transferred to a clean polyethylene vial and re-weighed. The sample was then analysed for elements producing radioactive half lives of 12.5–67 h (K, Na, As, Br, Au and Sb) by counting the γ rays emitted from the samples with a 12.7% relative efficiency Ge–Li detector coupled to a hardwired 4196 channel (Canberra) multichannel analyser. The Slowpoke

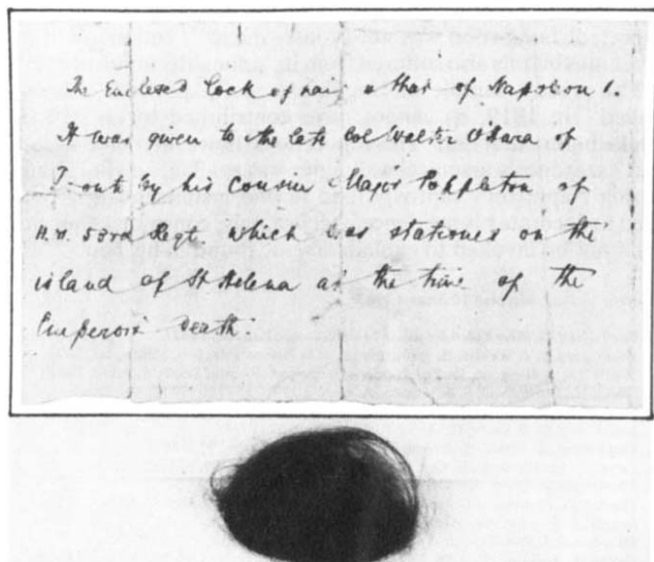


Fig. 1 Lock of hair from Napoleon I, with accompanying note.

reactor has the advantage of maintaining a reproducible and constant ($\pm 1\%$) neutron flux over the long term. The analytical system was calibrated previously with several standard reference materials, including NBS Coal (SRM 1632) and Fly Ash (SRM 1633), and with dilute solutions of arsenic salts. Stability of the specific activity produced by analysing various arsenic standards over nearly 6 yr guaranteed the reliability of the calibration.

The sample was counted for 5–50 min to obtain varying counting statistics. The concentrations of the elements K, Na, As, Br, Au and Sb in the hair were then calculated by comparing the net γ -ray peak areas with standard values. After preliminary counting, the hair was washed four times with absolute ethanol to remove surface contamination and was recounted several times over the next few days. The washing had no apparent effect on the elemental concentrations determined.

Table 1 shows that our sample of Napoleon's hair does not contain high levels of arsenic (maximum control range 0.7–1 p.p.m.), but does contain a moderately increased level of antimony (maximum control range 1–1.5 p.p.m.). Antimony (Sb) is of interest because it was a component of many medications of that time, particularly tartar emetic.

Previous investigations suggested that the maximal levels of arsenic were located 4 to 9 cm from the proximal end of a 13-cm length of hair². The average hair length of our sample, which measured 8 cm, was from the distal end, and therefore may not have recorded recent events. However, it would have contained high average levels of arsenic had there been chronic poisoning.

Improved technology in the past 20 years compared with that used in the original investigations^{2,4} allows us now to clearly separate the radioisotopes of arsenic, antimony and bromine. The previous analytical method used⁴ was not tested for the efficiency of separation of antimony and bromine from arsenic. Lack of such separation would tend to give apparent high

arsenic levels. In addition, our 12 mg hair sample from Napoleon could be a more representative sample, compared with the individual hair samples of less than 0.5 mg reported by Smith *et al.*²

We therefore conclude that Napoleon Bonaparte did not die of chronic arsenic poisoning.

Received 24 May; accepted 10 August 1982.

1. De Bourienne, L. A. F. *Mem. Napoleon Bonaparte* Vol 4, 427–428 (Grolier Society, London, 1885).
2. Smith, H., Forshufvud, S. & Wassen, A. *Nature* **194**, 725–726 (1962).
3. Weider, B. & Hapgood, D. *The Murder of Napoleon* (Methuen, New York, 1982).
4. Smith, H. *Analyt. Chem.* **31**, 1361–1363 (1959).
5. Takeuchi, T. *et al. Radioanalyt. Chem.* **70**, 29–55 (1982).

The formation of isomaltulose by immobilized *Erwinia rhapsontici*

Peter S. J. Cheetham, Carol E. Imber* & Jamie Isherwood*

Tate & Lyle Group Research & Development, Philip Lyle Memorial Research Laboratory, PO Box 68, Whiteknights, Reading RG6 2BX, UK

Recently the potential of immobilized cells for the production of biochemicals has become recognized^{1,2}. Here we describe a simple, stable, continuous method of converting concentrated sucrose solutions into isomaltulose using columns of *Erwinia rhapsontici* cells entrapped in alginate gel pellets, which exemplifies several strategies for maximizing the yield and stability of immobilized cells. This method has been successfully operated on a pilot-plant scale. The isomaltulose has possible applications, which are at present being evaluated by potential users, as a non-cariogenic bulking agent in foodstuffs and pharmaceuticals^{3–5}. Little enzyme activity or viability was lost during immobilization and surprisingly high yields of isomaltulose were obtained. The enzyme associated with the cells was stabilized by immobilizing in alginate rather than other support materials; by using structurally intact non-growing but not necessarily viable cells rather than isolated enzyme, disrupted cells or growing cell preparations; and most dramatically, by using concentrated pure sucrose as substrate and by maintaining complete conversion of the sucrose into isomaltulose. Thus the immobilized cells were about 350 times more stable than free cells, a half life of about 8,600 h being achieved.

Isomaltulose, commonly referred to as palatinose, is a reducing disaccharide (6-*O*- α -D-glucosylpyranosyl-D-fructofuranose) of about one-third the sweetness of sucrose but which otherwise has very similar physical and organoleptic properties to sucrose. It is believed to be non-cariogenic since very much less acid and glucan are formed by *Streptococcus mutans* from isomaltulose than from sucrose³. Several microorganisms have been reported to form isomaltulose^{6–10}; among them is *E. rhapsontici* (NCPB 1578), a cause of crown rot of rhubarb, which we have selected to use on the basis of the stability of its isomaltulose-forming enzyme when the cells are used in a non-growing immobilized form.

The adaptive significance of isomaltulose formation by *E. rhapsontici* is unclear. However, this bioconversion may be a method of irreversibly sequestering a carbon and energy source in a form less available to the host plant, and to other organisms that could otherwise compete successfully with the *Erwinia* cells for sucrose. Thus isomaltulose may have a similar useful role to the 3-ketosucrose formed by *Agrobacterium tumefaciens*^{11,12}, and the gluconate and 2-oxoglucose produced by *Pseudomonas aeruginosa*^{13,14}. Such a mechanism would be facilitated by the isomaltulose-forming enzyme being stabilized by sucrose, and by the ability of non-viable cells to continue

Table 1 Element concentration in Napoleon's hair

Element	Isotope	Half life (h)	γ -Ray energy (keV)	Concentration (p.p.m.)	
				Control ⁵	
K	⁴² K	12.5	1,225	750 $\pm$ 100	0.94–310
Na	²⁴ Na	15.0	1,369	1,100 $\pm$ 100	0.4–850
Br	⁸² Br	35.5	777	4.2 $\pm$ 0.6	0.77–490
As	⁷⁶ As	26.4	559	1.4 $\pm$ 0.2	0.079–0.67
Au	¹⁹⁸ Au	64.7	412	0.45 $\pm$ 0.06	0.0006–1.36
Sb	¹²² Sb	67.1	564	5.7 $\pm$ 0.6	0.016–1.30

Standard deviation, 2 σ (95% confidence level).

*Present addresses: Searle Pharmaceuticals (G. D. Searle & Co. Ltd), Lane End Road, High Wycombe, Bucks, UK (C.E.I.); Lavender Cottage, 1 Longreach, W. Horsley, Surrey, UK (J.I.).

to produce isomaltulose to the advantage of the remaining viable cells.

Cells were grown in an undefined medium and immobilized¹⁵⁻¹⁷ so as to enable continuous use of the enzyme and to prevent contamination of the product by the cells (Fig. 1). The enzyme associated with the cell is a previously undescribed sucrose-specific, direct glucosyl transferase which was most active at pH 7.0 and 30 °C, and had a K_m towards sucrose of 0.35 M. Activity was not enhanced by the presence of nitrogen, trace nutrients or other sugars. CO₂ and H₂, acids and a disaccharide believed to be 1-*O*- α -D-glucosylfructose⁸ were produced as side-products.

Immobilization in alginate stabilized the isomaltulose-synthesizing activity of intact cells and of a cell-free extract of the *Erwinia* cells (Table 1). Alginate was the most desirable support because in addition to its cheapness, lack of toxicity, mechanical strength and the simplicity of the immobilization method the cells lost only 5-10% of their activity on immobilization. Thus, alginate provides a favourable microenvironment for the cells, perhaps because it simulates the natural capsular polysaccharides produced by *Erwinia*, which are thought to have a protective role¹⁸. Of a variety of alternative immobilization techniques that we tested, none approached the combination of half life and efficiency of conversion achieved with alginate.

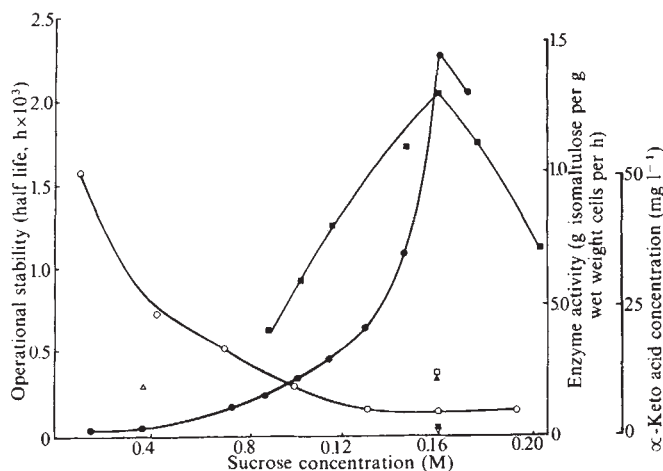


Fig. 1 The effect of sucrose concentration on the activities of free and immobilized *E. raphanici* cells and on the operational stability of the immobilized cells. *Erwinia* cells were grown, immobilized and used in columns as described in Table 1 legend. Immobilized cells (■) were shaken at 30 °C with sucrose solutions for initial rate determinations. For the determination of operational stabilities substrate solutions of varying sucrose content were pumped up columns using Watson-Marlow pumps such that an initial degree of conversion of sucrose to total products of 70% was achieved in each column (●) once a steady-state had been reached after 1 day's operation. Each column was run continuously for several months when necessary and half-life values calculated by linear regression analysis. The concentration of isomaltulose in the assay solutions or column eluates measured at 24-h intervals by assaying the reducing sugar concentration²³, or by HPLC using a Spherisorb 5 μ silica column which had been treated with acetone/water/tetraethylpentamine (80/20/0.01). This was also used as the eluting solvent. Sugars in the eluate were detected using a refractive index detector (Applied Chromatography Systems Ltd). Viable cells counts were made by grinding the pellets in phosphate-buffered saline (PBS), plating serial dilutions of the PBS extracts onto nutrient agar, incubating at 30 °C for 24 h and then counting the numbers of colonies formed. Total cell measurements were calculated from A_{540} readings of the PBS extracts made versus a suitable blank solution. α -Keto acids were measured as pyruvate equivalent by an adaptation of an existing method²⁴ (○-○). This method had an analytical threshold of 5 mg acids per l of column eluate. On one occasion affination syrup diluted to 1.6 M sucrose with deionized water and adjusted to pH 7.0 was supplied to a column of immobilized cells (□). On another 0.365 M sucrose pH 7.0 buffered with 100 mM HEPES buffer was supplied to a column (△). 76% and 95% conversion of the sucrose was achieved respectively. The continuous stirred reactor (▲) consisted of a stirred vessel containing 200 ml of pellets and 300 ml of substrate which was maintained at pH 7.0 by the addition of 0.1 M NaOH using a PHM64 pH meter and TTT60 titration unit (Radiometer). Substrate was pumped into the reactor at a rate of 0.016 l h⁻¹ and samples of the reactor contents pumped out at the same rate.

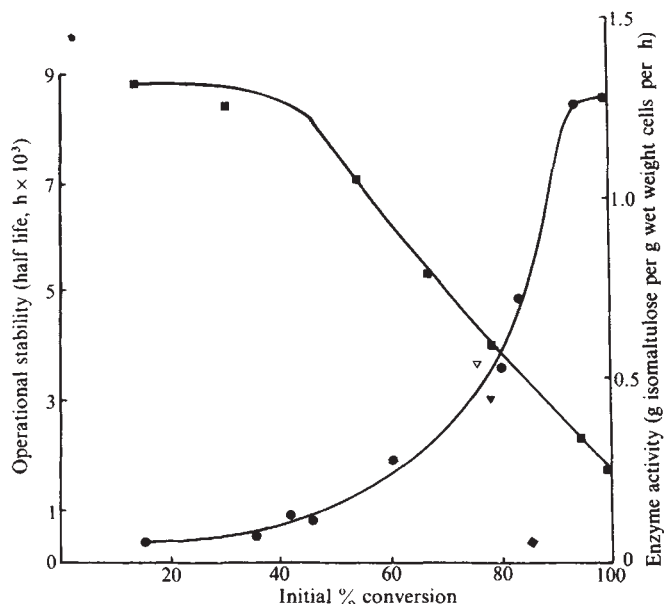


Fig. 2 The effect of the degree of conversion of sucrose into products on the operational stability of columns of immobilized *Erwinia* cells. Columns of immobilized cell pellets were prepared and operated as described in Table 1 and Fig. 1 legends with 1.6 M sucrose pH 7.0 pumped up the columns at various flow rates. The rate of decline in the activity of the immobilized cells was monitored while a constant degree of conversion was maintained by continuously decreasing the flow rate. The stability and activity of the immobilized cells are represented by ● and ■ respectively. Experiments were also carried out in the same manner in which benzylpenicillin (▼) or chloramphenicol (▽) were included in the substrate at concentrations of 500 μ g ml⁻¹.

The immobilized cells were both most active and most stable when high concentrations of sucrose were used, reaching a maximum at 1.6 M (Fig. 1). However, it was only the immobilized and not the free cells that were markedly stabilized by the concentrated sucrose solutions, although similar effects were observed when both high and low degrees of conversion of sucrose into isomaltulose were achieved (Table 1, Fig. 1). The use of concentrated sucrose was also advantageous because the equilibrium between isomaltulose and sucrose favoured the formation of isomaltulose. Also concentrated sucrose reduced the formation of gaseous and acidic side-products (Fig. 1), the growth of contaminant microorganisms and the activity of endogenous proteases, while minimizing the size of equipment and volumes of syrup that had to be manipulated, and the volume of water that had to be removed before crystallization.

The extent of conversion of sucrose into isomaltulose could be maximized (< 1% w/v sucrose remaining) by decreasing the flow-rate of substrate up the column, thus increasing its contact time with the cells. Increases in the extent of conversion were accompanied by decreases in activity as the reaction became first order with respect to the substrate concentration, and also by a marked increase in the stability of the immobilized cells (Figs 2, 3). This enhanced stability is probably due to diffusional restrictions on the immobilized enzyme activity especially as stability could be further increased by raising the concentration of entrapped cells.

High yields of isomaltulose were obtained first by screening for an organism resistant to excess substrate and product inhibition. Second, a concentrated substrate lacking nutrients was used, such that substrate was not diverted into cell growth, and other enzymes of the cells that tend to produce side-products such as acids (Fig. 1) are inhibited. Third, the use of packed-bed reactors, in which virtual plug-flow of substrate is achieved by using even-sized, spherical, evenly packed immobilized cell pellets¹⁶, encouraged complete conversion of substrate, plug-flow being impossible with free cells or stirred reactors¹⁹.

In the above conditions the immobilized cell columns yielded around 0.75 g isomaltulose per g of sucrose supplied and had an initial activity of about 0.2 g isomaltulose produced per g

wet weight of cells per h, equivalent to 40 g isomaltulose produced per l column volume per h. The activity decayed very slowly with a half life of 1 yr (Fig. 2, Table 1). Small, white, needle-shaped, but virtually pure crystals of isomaltulose could be recovered by evaporating the column eluate under vacuum at 60 °C until an isomaltulose concentration of 2.0 M was reached, seeding the supersaturated solution as it cooled and collecting the resulting crystals in a basket centrifuge. This purification and recovery of the crystals was greatly facilitated by the use of concentrated sucrose solutions as substrate and by the achievement of high degrees of conversion by the cells.

Rapid loss of the enzyme activities of non-growing microorganisms often limits their useful life as biosynthetic agents, particularly of secondary metabolites. Enzymes vary considerably in stability and very little is understood of the causes of inactivation and of the optimum physiological state of cells when they are to be used in a non-growing form. Nevertheless the stability of the *Erwinia* cells used here is surprisingly high considering that the normal metabolic turnover of enzymes usually rises substantially when growth ceases^{20,21}, and some enzymes disappear at rates greater than the overall rate of protein turnover²². Stabilization is not due to product stabilization by the isomaltulose (Table 1), but may be connected with the cells' low rate of endogenous metabolism as illustrated by the decreased rate of acid formation (Fig. 1), or with their plasmolysis by the high sugar concentrations. The purity of the sucrose used was also important, as both the activity and stabil-

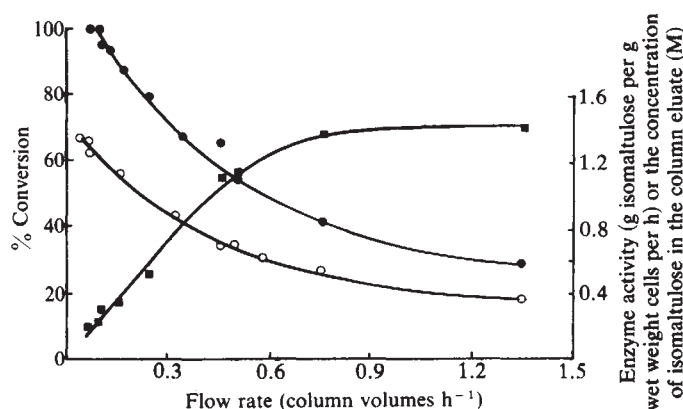


Fig. 3 The influence of the residence time of substrate in the column on the activity of the immobilized *Erwinia* cells (●—●) and the degree of conversion of sucrose to products (○—○). *E. rhapsodica* were grown and immobilized, and the columns operated as described in Table 1 legend. The flow rate through the column is expressed as the number of empty column volumes (ecv) of substrate pumped through the column per h where an empty column volume represents the volume of the pellets plus the interstitial fluid. Thus a flow rate of 0.08 ecv h⁻¹ through the 0.035–1 columns used in these experiments gives a residence time of 12.5 h. The molar isomaltulose concentration in the column eluate is expressed as ○—○.

Table 1 The operational stabilities of the isomaltulose forming activities of various *E. rhapsodica* preparations

Type of preparation and conditions of use	Stability (half life, h)
Free cells maintained in exhausted growth medium	335
Free cells used batchwise in 0.12 M sucrose	25
Free cells used batchwise in 1.60 M sucrose (28% conversion) [†]	41
Immobilized cells used with 0.12 M sucrose (70% conversion)	55
Immobilized cells used with 1.60 M sucrose (70% conversion)	2,265
Immobilized cells used with 1.60 M sucrose (99% conversion)	8,625*
Immobilized cell debris (45% conversion)	190
A cell-free extract obtained by osmotically shocking cells used batchwise	26
An immobilized cell-free extract (81.5% conversion)	620
Immobilized cells supplied continuously with growth medium (90% conversion)	76
Immobilized cells supplied with 1.02 M isomaltulose	300

Erwinia rhapsodica (NCPBP 1578, ATCC 29283) a Gram-negative facultative anaerobe was grown in shake flasks at 30 °C containing 200-ml aliquots of medium consisting of sucrose (40 g l⁻¹), peptone (10 g l⁻¹) and beef extract (4 g l⁻¹). Once stationary phase had been reached after about 70 h the cells were collected by centrifuging at 23,000g for 20 min at 30 °C. A yield of approximately 7 g wet packed cells per l of medium was obtained (~2 × 10⁹ viable cells per g wet weight). The cells were evenly dispersed as a 20% (wet w cell/v) slurry in a freshly prepared 5% (dry w/v) sodium alginate solution (pH 6.0), which was then immobilized by extruding dropwise, gelation taking place at room temperature in a 0.1 M CaCl₂ solution (pH 6.5) to form even-sized, smooth, spherical pellets with a diameter of 5 mm and consisting of a thin, highly polymerized skin and a core of highly porous gel. Sodium alginate (Protanal LF10/60, BDH) extracted from *Laminaria hyperborea*, which has a high guluronic to mannuronic acid ratio and thus a high mechanical strength, was used in all the experiments, the gel being formed by calcium ions cross-linking guluronic acid residues in adjacent polysaccharide chains. The pellets were packed into columns (20 × 1.5 cm diameter) thermostated at 30 °C and supplied with a sucrose solution adjusted to pH 7.0 with 0.1 M NaOH. The enzymatically active cell-free extract was prepared by suspending 1.5 g (wet weight) cells in 150 ml of ice-cold deionized water for 30 min, followed by centrifugation at 23,000g for 20 min at 30 °C, so as to obtain a clear dilute enzyme solution. The cell-free extract could not be easily concentrated by ultrafiltration but could be permanently immobilized by mixing with 8% w/v alginate and gelling with 1.0 M CaCl₂. *Erwinia* cell debris was prepared by mixing 20 g (wet weight) cells with 20 g dry sand and 20 ml of deionized water and shaking for 20 min at 20 °C in a Mickle-shaker (Chem. Lab. Instruments), so that maximum oscillation of the vials occurred. The growing cell preparation was prepared by immobilizing a 1% (wet weight/v) cell inoculum and pumping medium through the column continuously at a rate of 0.04 empty column volumes per h. Unless otherwise stated preparations were assayed in columns at 30 °C using 1.60 M sucrose pH 7.0 as substrate.

*Stability is more properly quoted as a first order decay constant of 1.95×10^{-2} per day.

[†]Because of the interdependence of stability and degree of conversion (Fig. 2) where appropriate the degree of conversion is quoted in parentheses above.

ity were greatly reduced when affination syrup, an impure sucrose stream produced during refining, was used (Fig. 1). Loss of activity was not due to the action of contaminant microorganisms, and no enzyme induction or turnover could have occurred in the immobilized *Erwinia* cells considering the purity of the substrate used and as addition of chloramphenicol or benzylpenicillin to the substrate, which inhibit protein and cell wall synthesis respectively, did not affect stability (Fig. 2). Stability was, however, enhanced by the use of buffered substrate (Fig. 1) and by the use of the immobilized cells in packed columns rather than in a continuous stirred reactor maintained at pH 7.0 (Fig. 1), or a fluidized-bed in which the reactants were recycled through the reactor and neutralized after each cycle.

The stability of the immobilized cells was much greater than when a cell-free enzyme extract, mechanically disrupted cells or growing cells were used (Table 1), although the initial activities of these preparations were not appreciably different. Viable cells were not required for isomaltulose formation because whereas the activity of the cells decayed slowly and linearly with time, with a half life of 1 yr (Table 1, Fig. 2), the viable cell count declined rapidly with a half life of 300 h. Some metabolic activity was, however, retained by the non-viable immobilized cells as they continued to produce α -keto acids throughout the useful life of the column, presumably by the mixed acid and 2,3-butanediol pathways, such that typically the column eluate was at pH 4.0–4.5.

In conclusion, we have succeeded in obtaining high yields of isomaltulose, stabilizing the isomaltulose synthesizing activity of the *Erwinia* cells and operating the process on a small pilot-plant scale. The immobilized cells appear to be in a resting state caused by a lack of nitrogenous and trace nutrients and the physical constraints imposed by the alginate gel as they did not grow or autolyse and no involution forms were observed. These experiments demonstrate the potential of immobilized non-growing cells particularly when the enzyme of interest is designed to function when the cells are in a non-growing form. They also illustrate the desirability of using specific, stable catalysts, mild immobilization techniques, high reactant concentrations and of maintaining high degrees of conversion and of developing efficient methods for purifying and recovering the product.

This work is the subject of UK Patent Application GB 2 063 268A.

Received 13 May; accepted 2 August 1982.

- Abbott, B. J. A. *Rep. Ferm. Processes*, 205–235 (1977).
- Cheetham, P. S. J. *Topics Enzym. Ferm. Biotechnol.* 4, 189–238 (1980).

3. Roberts, K. R. & Hayes, M. C. *Scand. J. dent. Res.* **88**, 201–209 (1980).
4. UK Patent Application 2 066 639A assigned to Tate & Lyle PLC.
5. UK Patent Application 2 066 640A assigned to Tate & Lyle PLC.
6. Sharpe, E. S., Stodola, F. H. & Koepsell, H. J. *Am. chem. Soc. Abstr.*, 5-D (1954).
7. Bourne, E. J., Hutson, D. H. & Weigel, H. *Biochem. J.* **79**, 549–553 (1961).
8. Lund, B. M. & Wyatt, G. M. *J. gen. Microbiol.* **78**, 331–336 (1973).
9. Weidenhagen, R. *Zucker* **14**, 456–462 (1961).
10. UK Patent Specification 1 429 334 assigned to the South German Sugar Co.
11. Bernaerts, M. J. & De Ley, J. *Biochim. biophys. Acta* **30**, 661–662 (1958).
12. Hayano, K. & Fukui, S. *J. biol. Chem.* **242**, 3665–3672 (1967).
13. Whiting, P. H., Midgley, M. & Dawes, E. A. *Biochem. J.* **154**, 659–668 (1976).
14. Whiting, P. H., Midgley, M. & Dawes, E. A. *J. gen. Microbiol.* **92**, 304–310 (1976).
15. Kierstan, M. & Bucke, C. *Biotechnol. Bioengng* **19**, 387–397 (1977).
16. Cheetham, P. S. J., Blunt, K. W. & Bucke, C. *Biotechnol. Bioengng* **21**, 2155–2168 (1979).
17. Cheetham, P. S. J. *Enzym. microb. Tech.* **1**, 183–188 (1979).
18. Politis, D. J. & Goodman, R. N. *Appl. env. Microbiol.* **40**, 596–607 (1980).
19. Lilly, M. D. & Sharpe, A. K. *Chem. Engng. Lond.* **215**, CE12 (1968).
20. Mandelstam, J. *Bact. Rev.* **24**, 289–308 (1960).
21. Trinci, A. P. J. & Thurston, C. F. *Symp. Soc. gen. Microbiol.* **26**, 55–80 (1976).
22. Thurston, C. F. *Process Biochem.*, **7**, (8) 18–23 (1972).
23. Astoor, A. & King, E. J. *Biochem. J.* **56**, XLIV (1954).
24. Sloneker, J. H. & Orentos, D. G. *Nature* **194**, 478–479 (1962).

Changes in muscle stiffness during contraction recorded using ultrasonic waves

Yojiro Tamura, Ichiro Hatta & Takashi Matsuda

Department of Applied Physics, Faculty of Engineering,
Nagoya University, Nagoya 464, Japan

Haruo Sugi* & Teizo Tsuchiya

Department of Physiology, School of Medicine, Teikyo University,
Tokyo 173, Japan

Although it is generally accepted that muscles contract by means of cross-links between the thick and thin filaments^{1,2}, the molecular mechanism of contraction is still a matter for debate and speculation. To investigate the number and state of the cross-links at various stages of muscle contraction, muscle stiffness changes have been studied by applying step or sinusoidal length changes and measuring the resulting force changes^{3–5}, or the propagation of longitudinal mechanical waves^{6,7}. These methods, however, involve relatively large perturbations to the contractile system, and may not be free from the possibility that the state of the contractile system is altered by the measurement procedure. We have developed a technique in which muscle stiffness can be continuously recorded during a single mechanical response in frog skeletal muscle by measuring the propagation velocity of ultrasonic waves with negligibly small perturbations to the contractile system and a high time resolution. Using this technique we have obtained the novel and unexpected result that during contraction muscle stiffness decreased in the transverse direction, though it increased in the longitudinal direction.

For measuring muscle stiffness in the longitudinal direction, the semitendinosus muscle of the bullfrog (*Rana catesbeiana*) was mounted in a Lucite chamber as shown in Fig. 1a; a ceramic piezo-electric transducer, which was a circular plate (8 mm diameter) of PbZrO₃–PbTiO₃ (ref. 8), was in contact with the muscle at each bent end portion. One transducer was repetitively energized with sinusoidal voltages from a function generator (Wavetek, model 162) to transmit brief trains of ultrasonic waves (1, 3 and 7 MHz) to the muscle. The wave trains propagated longitudinally through the muscle, and were sensed by another transducer.

For measuring muscle stiffness in the transverse direction, the sartorius muscle of the bullfrog was mounted with its lower surface in contact with the bottom of the chamber (Fig. 1b); a ceramic transducer (diameter, 6 mm) was in contact with the upper muscle surface. The ultrasonic wave trains (7 MHz) were transmitted to the muscle by the transducer, propagated across

the muscle, and reflected by the chamber to be sensed, as trains of echo waves, by the transmitting transducer.

For accurate determination of the change in propagation velocity of ultrasonic waves in the muscle, the time interval between the original and the propagated ultrasonic signals was divided into two parts, T and ΔT (Fig. 2). A rectangular pulse of a constant duration T with an accuracy of 9.5 figures was generated by a synthesizer (Rockland, type 5100). The duration ΔT was equal to the time interval between the end of the pulse T and the beginning of the first propagated signals. The pulse with a duration ΔT was integrated by a time-to-amplitude converter to obtain an amplitude signal A proportional to the time ΔT . As the propagation velocity of ultrasonic waves changed with time to give values $\Delta T_1, \Delta T_2, \dots$, the corresponding amplitude signals $A_1, A_2, \dots$, were successively recorded in a digital-memoryscope (Hitachi, type VC 801). The time resolution of muscle stiffness measurements was limited by the frequency of repetition of ultrasonic wave trains, being 250 μ s for the longitudinal stiffness and 30 μ s for the transverse stiffness. The amplitude of perturbations produced in the muscle during the application of ultrasonic waves (~ 1 Å) was negligibly small. The experiments were performed at room temperature (20–25 °C).

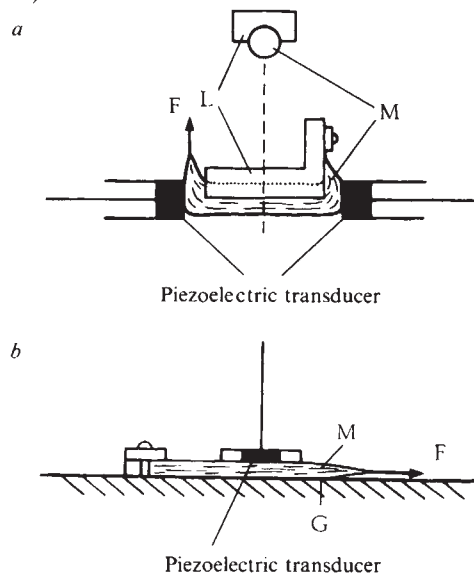


Fig. 1 Methods of recording muscle stiffness changes during isometric contraction. *a*, Experimental arrangement for measuring longitudinal stiffness. The semitendinosus muscle was held by a Lucite block (L) with a cylindrical groove of ~ 5 mm diameter, and bent around the rectangular corner of the block near both ends. Two ceramic transducers were then put in contact with the muscle in such a way that their planes were perpendicular to muscle long axis. The distance between the transducers was 2–2.5 cm. One transducer transmitted ultrasonic waves to the muscle, while the other sensed the propagated waves. *b*, Arrangement for measuring transverse stiffness. The sartorius muscle was placed with its lower surface in contact with the bottom of the chamber (G) serving as a reflecting plane of ultrasonic waves. The transducer was in contact with the upper muscle surface so that its plane was in parallel with the bottom of the chamber. The distance between the transducer and the chamber was 1.5–2 mm. Ultrasonic waves from the transducer propagated across the muscle, and were reflected by the bottom of the chamber to be sensed by the same transducer. In both *a* and *b*, one end of the muscle was fixed in position, while the other end was connected to the strain gauge (Sinko, type UT, F). The muscle was soaked in Ringer solution (NaCl 115 mM, KCl 2.5 mM, CaCl₂, 1.8 mM, pH 7.3 by Na-phosphate buffer), and each time of stiffness measurement the chamber was drained and the muscle was stimulated maximally with single or repetitive 0.5-ms pulses through a pair of Pt plate electrodes in contact with the muscle surface along its entire length (not shown). To minimize the muscle movement, the initial muscle length was made at 1.2 times slack length. The changes in propagation velocity and isometric force were recorded in a memoryscope, and displayed on a chart recorder.

* To whom correspondence should be addressed.

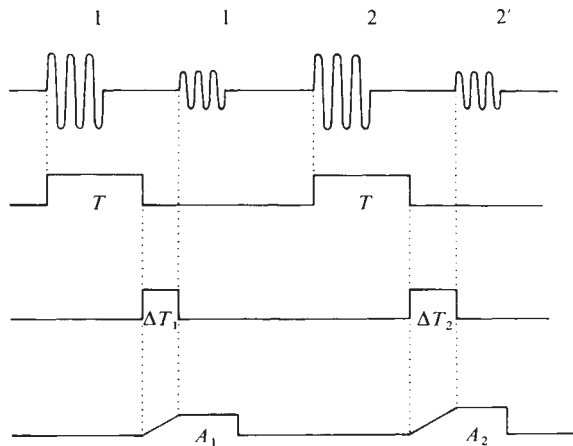


Fig. 2 Diagram illustrating the method for measuring propagation velocity changes. Top record shows two original wave trains (1 and 2) and the corresponding propagated wave signals (1' and 2'). As shown in the two middle records, the interval between the original and the propagated signals was divided into two parts; a rectangular pulse of constant duration (T) and an additional brief pulse of variable duration ΔT . The brief additional pulses (ΔT_1 and ΔT_2) were integrated by the time-to-amplitude converter to give the amplitude signals (A_1 and A_2) in the bottom record.

Figure 3a shows a typical example of the change in propagation velocity of ultrasonic waves along the muscle length during an isometric twitch. On stimulation, the propagation velocity started to increase before the onset of isometric force development, reached a maximum while the isometric force still continued to rise, and then began to decrease to the initial value. The return of propagation velocity to the initial value was complete before the completion of relaxation of isometric twitch. With 3 and 7 MHz waves, the longitudinal propagation velocity was the same within the accuracy of measurements, being $1.61 \pm 0.05 \times 10^3 \text{ m s}^{-1}$ (mean $\pm$ s.d., $n = 20$) in resting muscle. The magnitude of increase in the propagation velocity during twitches was always about 0.9%. Smaller values of propagation velocity was obtained with 1 MHz waves.

A typical example showing the change in propagation velocity in the direction perpendicular to the muscle longitudinal axis during a twitch is shown in Fig. 4a. Unexpectedly, the transverse propagation velocity was found to decrease during a twitch. With 7 MHz waves, the transverse propagation velocity was $1.50 \pm 0.1 \times 10^3 \text{ m s}^{-1}$ (mean $\pm$ s.d., $n = 20$) in resting muscle, and always decreased by about 1% during a twitch. The propagation velocity started to decrease before the onset of isometric force development, and reached a minimum prior to the peak twitch force (Fig. 3a).

In both the longitudinal and transverse propagation velocity, a steady propagation velocity was attained during the plateau of an isometric tetanus. The magnitude of velocity changes during a tetanus was about twice that during a twitch. When the peak twitch force was varied by applying submaximal stimuli of various strengths, the magnitude of both the longitudinal and transverse velocity changes was proportional to the peak twitch force (Figs 3b, 4b).

When mechanical waves with a constant frequency f propagate in muscle, their phase velocity V_f is related to muscle stiffness E_f as:

$$V_f = (E_f/\rho)^{1/2} \quad (1)$$

where ρ is the density of muscle, which may remain unchanged during isometric contraction. Thus, muscle stiffness can be estimated from equation (1) if the measured velocities, which are generally very close to the group velocities, can be regarded as the phase velocities. This condition is satisfied in the range of high enough wave frequencies where the frequency dependence of the propagation velocities is no longer obvious. The similarity of the propagation velocity between 3 and 7 MHz

waves indicates that the measured velocities with waves above 3 MHz are equal to the phase velocities. Moreover, in the case of the longitudinal velocities, the wavelength λ should also satisfy the condition, $a/\lambda > 2.5$, where a is the muscle radius⁹; otherwise, the waves propagate along the muscle fibres with mixed modes to make the measurement of the phase velocity complicated. In the present study, the muscle was nearly circular in cross-section (Fig. 1) with a radius of about 2.5 mm, while the wavelengths of 7, 3 and 1 MHz waves were about 0.2, 0.5 and 1.5 mm respectively.

Therefore, the velocities with waves above 3 MHz can serve as a valid measure of muscle stiffness, and the results shown in Figs 2 and 3 may indicate that, during isometric contraction, longitudinal muscle stiffness increases while transverse muscle stiffness decreases. The linear relation between the magnitude of velocity changes and the peak twitch force may be taken to indicate that the stiffness changes observed are related to the force generating mechanism. Note also that the time course of stiffness changes during the isometric force development resembles that of the intensity ratio of two equatorial reflections $(I_{1,0}/I_{1,1})^{10,11}$, which is regarded to reflect the radial movement of cross-bridges towards the thin filaments.

Assuming that ρ is 1 g cm^{-3} , the value of longitudinal muscle stiffness for 3 MHz waves was estimated from equation (1) to be about $2.56 \times 10^9 \text{ N m}^{-2}$ in resting muscle, and the increment of muscle stiffness in a tetanus above the resting value was about $6 \times 10^7 \text{ N m}^{-2}$. The high value of resting muscle stiffness results from the fact that, with ultrasonic waves, it is the bulk properties of the muscle that are measured, and therefore it is the increment of stiffness during contraction that is relevant to the cross-bridge properties. We tentatively calculated the increment of muscle stiffness for 3 kHz mechanical waves during contraction from the published data of Truong⁷, and also obtained a value of the order of 10^7 N m^{-2} . A similar value of muscle stiffness is also obtained from the slope of the T_1 curve in the length step experiments of Ford *et al.*³. This suggests

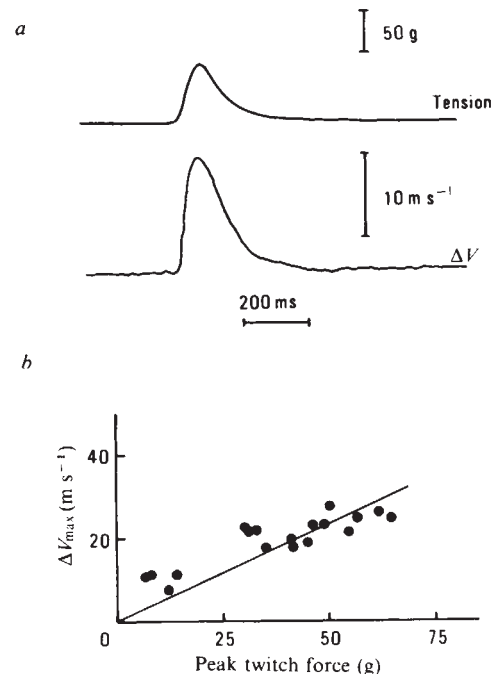


Fig. 3 a, Change in longitudinal propagation velocity of 3 MHz ultrasonic waves (ΔV , lower trace) during an isometric twitch (upper trace). In this and subsequent figures, the upward and downward deflection of the velocity trace indicate an increase and a decrease in propagation velocity respectively. Temperature, 25 °C. b, Relationship between the magnitude of increase in the longitudinal propagation velocity in a twitch (ΔV_{max} , ordinate) and the peak twitch force (abscissa). Peak twitch forces were varied by applying submaximal stimuli of various strengths. Temperature, 25 °C.

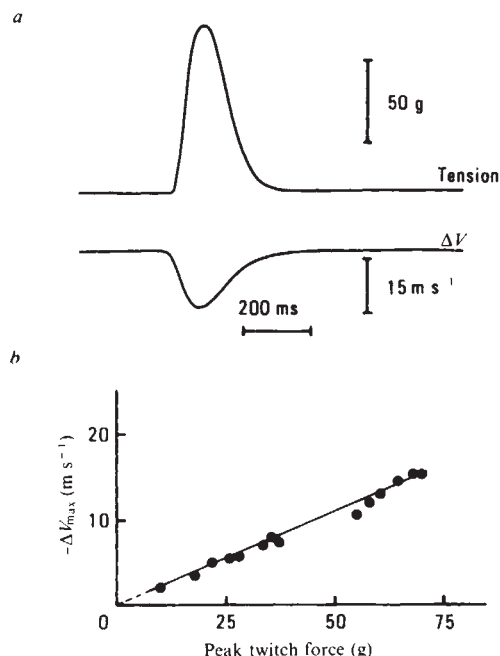


Fig. 4 *a*, Change in transverse propagation velocity (ΔV) of 7 MHz waves during an isometric twitch. Note that the propagation velocity decreases during a twitch. Temperature, 15 °C. *b*, Relation between the magnitude of decrease in the transverse propagation velocity in a twitch ($-\Delta V_{\max}$, ordinate) and the peak twitch force (abscissa). Peak twitch forces were varied by applying submaximal stimuli of various strengths. Temperature, 25 °C.

that the increase in longitudinal muscle stiffness during contraction, which has been observed with different techniques, may result from the formation of an elastic element exhibiting almost constant stiffness values over a wide frequency range of mechanical perturbations. On the other hand, transverse muscle stiffness for 7 MHz waves was about $2.25 \times 10^9 \text{ N m}^{-2}$ in resting muscle, and the decrement of muscle stiffness in a tetanus was about $8 \times 10^7 \text{ N m}^{-2}$. The decrease in transverse stiffness may not result from local segmental shortening introducing an increase in protein density underneath the transducer, as the stiffness decrease was always observed irrespective of the position of the transducer along the entire muscle length. Moreover, our theoretical calculations indicate that an increase in protein density leads to an increase in propagation velocity, an effect opposite to the present results.

A possible mechanism of the decrease in transverse stiffness, which is a novel and unexpected result of the present study, may be that the nature of the cross-links between the filaments is such that the strain in the transverse direction is largely taken up by the deformation of the cross-links to result in softening of the muscle in the transverse direction.

We thank Professor Sir Andrew F. Huxley for his invaluable advice and criticisms during this work.

Received 24 May; accepted 9 August 1982.

- Huxley, A. F. *Prog. Biophys. biophys. Chem.* **7**, 255–318 (1957).
- Huxley, H. E. in *The Cell* (eds Brachet, J. & Mirsky, A. E.) 365–481 (Academic, New York, 1960).
- Ford, L. E., Huxley, A. F. & Simmons, R. M. *J. Physiol., Lond.* **269**, 441–515 (1977).
- Rüegg, J. C. et al. in *Cross-bridge Mechanism in Muscle Contraction* (eds Sugi, H. & Pollack, G. H.) 125–148 (University of Tokyo Press and University Park Press, Baltimore, 1979).
- Julian, F. J. & Sollins, M. R. *J. gen. Physiol.* **66**, 287–302 (1975).
- Schoenberg, M., Wells, J. B. & Podolsky, R. J. *J. gen. Physiol.* **64**, 623–642 (1974).
- Truong, X. T. *Am. J. Physiol.* **226**, 256–264 (1974).
- Berlincourt, D. A., Curran, D. R. & Jaffe, H. in *Physical Acoustics* Vol. 1, pt A (ed. Mason, W. P.) 169–270 (Academic, New York, 1964).
- Yu, L. T., Brennan, J. N. & Sauer, J. A. *J. Acoust. Soc. Am.* **27**, 550–555 (1955).
- Haselgrove, J. C. & Huxley, H. E. *J. molec. Biol.* **77**, 549–568 (1973).
- Amemiya, Y., Sugi, H. & Hashizume, H. in *Cross-bridge Mechanism in Muscle Contraction* (eds Sugi, H. & Pollack, G. H.) 425–443 (University of Tokyo Press and University Park Press, Baltimore, 1979).

Induction of immortality is an early event in malignant transformation of mammalian cells by carcinogens

Robert F. Newbold, Robert W. Overell
& Jane R. Connell

Chemical Carcinogenesis Division, Institute of Cancer Research,
Pollards Wood Research Station, Nightingales Lane,
Chalfont St Giles, Buckinghamshire HP8 4SP, UK

Limited lifespan in normal diploid mammalian cells in culture is well documented^{1–3} and there is evidence that it is also a property of normal tissues serially transplanted *in vivo*^{4,5}. In contrast, in favourable growth conditions, cells derived from tumours or premalignant tissue often exhibit unlimited proliferative potential (immortality) both *in vitro* and *in vivo*^{2,4,6–10}. Malignant transformation of diploid cells in culture by carcinogens has been clearly demonstrated in several laboratories^{2,11–13}. However, although it has been frequently reported that establishment in culture is one of many alterations in cell behaviour associated with transformation, the induction of immortality and its role in malignancy have not previously been subjected to serious scientific examination. Here, we show that treatment of Syrian hamster cell cultures with carcinogenic chemicals can induce rare immortal variants, the progeny of which then frequently progress to anchorage independence and malignancy after a further period of growth. In addition, by studying the properties of cell lines that are either immortal or anchorage-independent but not both, we have obtained evidence that immortality, although insufficient by itself, may be a prerequisite for malignant transformation.

Rapidly dividing fusiform cells which, in minimal media, dominate cell cultures derived from primary tissue, are generally assumed to be fibroblasts. Even though the precise origin of cells which adapt easily to culture has not been established¹⁴, they have formed the basis of most cell culture transformation systems developed to date. Nevertheless, their use for this purpose can be justified by the fact that malignant tumours can arise *in vivo* from most cells that are capable of division, including those of mesenchymal origin. The cells used in our experiments were obtained from mixed Syrian hamster embryo tissues, newborn hamster dermis and human fetal lung; they were maintained in Dulbecco's modified Eagle's minimal essential medium (DMEM; Gibco) supplemented with 15% fetal calf serum (FCS). Spontaneous malignant transformation is rare in cells derived from Syrian hamster^{11,13} and has never been reported in the case of human fibroblasts from normal donors¹⁵.

The effect of several carcinogenic chemicals on the frequency of appearance of immortal Syrian hamster cell variants is shown in Table 1. Replicate cultures of tertiary embryo or dermal cells were treated with solvent alone (dimethyl sulphoxide, DMSO) or solutions of carcinogens (see Table 1). In controls and some treated cultures a phase of rapid proliferation, necessitating frequent subculture, was followed by a decline in growth rate, leading eventually to a complete loss of division potential. At this point, the cells became greatly enlarged and flattened, and phase-dense stress fibres¹⁶ in the cytoplasm were clearly visible by phase contrast microscopy. These stages in culture lifespan are typical of those described for human fibroblasts by Hayflick and Moorhead¹; the term senescence has since been widely used to denote the final non-proliferative state. In a proportion of cultures treated with benzo(a)pyrene (BP), its ultimate carcinogenic metabolite *anti*-BP-7,8-dihydrodiol-9,10 oxide (BPDE)¹⁷, and the two methylating agents *N*-methyl-*N*-nitrosourea (MNUA) and dimethylsulphate (DMS)¹⁸, immortal (Sen⁺) cells appeared which gave rise to continuous cell lines. (Treated cells capable of proliferation for

Table 1 Induction of immortal variants in Syrian hamster cell cultures treated with various chemical carcinogens

Expt	Treatment	Dose ($\mu\text{g ml}^{-1}$ medium)	Survival (% control)	No. of immortal cultures obtained/ no. of replicate cultures observed
<i>a, Mixed embryo cultures</i>				
1	BP	1.0	90	8/18
2	BPDE	0.1	81	1/7
	BPDE	0.3	20	2/4
	DMS	12.5	84	1/6
	DMS	25.0	21	2/6
3	BP	1.2	96	18/21
	BPDE	0.2	56	8/19
	DMS	19.0	38	3/21
	MNUA	34.0	73	9/20
Pooled control	DMSO (0.2%)		100	1/101
<i>b, Newborn dermal fibroblasts</i>				
1	BP	1.0	100	3/10
	BPDE	0.06	95	1/10
	BPDE	0.12	82	3/10
2	BP	2.5	95	4/10
	BPDE	0.07	94	2/9
	BPDE	0.15	63	2/9
	MNUA	30.0	60	2/9
Pooled control	DMSO (0.2%)		100	0/50

Primary cell cultures were prepared from trypsinized, minced, eviscerated Syrian hamster embryos (13–14 days of gestation) and collagenase-treated newborn dermis. In the latter case, separation of the epidermis from the dermis was facilitated by floating skins overnight in 0.01% pronase (Sigma) at 4°C as described elsewhere²⁸. Disaggregated cells of both types were seeded at a density of 1.5×10^7 per 175-cm² culture flask (Nunc) and, when almost confluent, were trypsinized and cryopreserved in liquid nitrogen. Samples of cells from each preparation were checked for the presence of mycoplasma by staining with the fluorescent dye, Hoechst 33258. For experiments, ampoules were thawed and the cells returned to 175-cm² flasks. After 3–4 days at 37°C, exponentially growing cultures were trypsinized and 1×10^6 cells seeded in 75-cm² flasks in 20 ml of medium for treatment with carcinogens. Compounds were added directly to the flasks 24 h later as 500× stock solutions in DMSO. Cells were exposed to the agents for 1 h (MNUA, DMS and BPDE) or 48 h (BP) after which the medium was changed. Embryo cells were re-plated 2 h later in replicate 5-cm Petri dishes (Sterilin) at a density of 0.4×10^5 per dish; dermal cells were allowed to grow to confluence. From this point onwards, cells were subcultured when confluent (embryo cells 1:10 in 5-cm dishes, dermal fibroblasts 1:5 in 9-cm dishes) and the medium replaced regularly every 3 days. Cell survival was determined immediately after treatment by plating 400 cells from each experimental group into 5-cm dishes on a feeder layer of 8×10^4 lethally irradiated homologous cells; colonies were scored after 7–10 days of incubation. Control plating efficiencies ranged from 22 to 34% for both cell types, depending on the cell preparation and the serum batch used. On reaching senescence, mass cultures were subcultured once more if confluent (1:2). The resulting semi-confluent cultures of non-proliferating cells remained visibly healthy for many weeks afterwards, if the medium was replaced regularly. The carcinogen treatments described had no significant effects on the replicative lifespan of cultures which showed no transformation. Transformed cell lines were maintained in culture for at least 100 pd after their emergence and cells were regularly cryopreserved for subsequent studies of their tumorigenicity and anchorage independence.

>100 population doublings (pd) after senescence of the respective controls, were considered immortal.) Transformation or senescence of cultures was found to be highly reproducible when experiments were repeated with replicate samples of cells cryopreserved soon (6–8 pd) after treatment, demonstrating that the system permits efficient recovery of transformants. On the assumption that Sen⁻ variants were monoclonal in origin, we estimated the range of transformation frequencies for the embryo and dermal cells to be $0.9\text{--}7 \times 10^{-5}$ and $1.6\text{--}6 \times 10^{-7}$, respectively, per viable treated cell, depending on the agent and the dose applied. More detailed, quantitative studies are now in progress to determine accurately the relative potencies of these and other carcinogens and mutagens for inducing immortality. We are also examining the validity of the assumption that Sen⁻ lines which emerge in these conditions are clones, by performing experiments with mixtures of male and female cells of known transformability and using C-banding to reveal the presence of the heterochromatic Y-chromosome in transformed cells.

With hamster embryo cells, the appearance of carcinogen-induced Sen⁻ variants occasionally occurred after crisis at the end of the proliferative phase of growth but more often little

or no crisis was observed, indicating that immortal cells had emerged before the cessation of normal cell division. In contrast, Sen⁻ dermal fibroblasts always appeared at the end of the normal culture lifespan as foci of dividing cells on a background of senescent counterparts. Only one Sen⁻ cell line was obtained from over 100 cultures of solvent-treated control embryo cells and this line (S33) has generated no malignant progeny on further subculture; no spontaneous transformation has so far been observed with the dermal cells. Because of their shorter lifespan (15–20 pd compared with 50–60 pd for the hamster embryo cells) and the fact that Sen⁻ transformants can be scored as foci 6–7 weeks after treatment, the dermal cells seem preferable for future studies. When foci of Sen⁻ dermal cells were isolated and grown in mass culture, it was noted that in many lines enlarged, flattened, non-proliferating cells continued to be produced at a variable rate. Indeed, in a few lines, the probability of senescence was so high that their continued maintenance *in vitro* required careful handling at subculture. In contrast, some cell lines were obtained which grew almost exponentially with few senescent cells visible. Thus, the induction of immortality did not seem to be an all-or-none process; rather it appeared to involve different degrees of reduction in the commitment of cells to a non-proliferative state.

When injected subcutaneously or intraperitoneally into syngeneic 1-day-old hamsters, carcinogen-induced Sen⁻ cell lines were not tumorigenic immediately they exceeded the normal lifespan but most (75%) progressed to malignancy after a further period of culture. This period was highly variable and was a characteristic of the individual cell line; some lines remained non-tumorigenic even after prolonged growth (>300 pd) *in vitro*. In many cases, the tumours formed were capable of invasion of surrounding tissues and/or metastasis to other organs. Normal untreated cells did not give rise to tumours in any circumstances. As reported previously¹³, the acquisition of tumorigenic potential was correlated perfectly with the

Table 2 Correlation between anchorage-independent growth and tumorigenicity in immortal Syrian hamster cells

Cell line	Passage tested (approximate pd)	Frequency of Aga ⁺ colonies per seeded cell	No. of animals with tumours/no. surviving	Mean latent period for tumour formation (days)	No. of animals with metastases/no. with tumours
S33	48 (190)	<10 ⁻⁶	0/5	—	—
BP1	54 (215)	<10 ⁻⁶	0/9	—	—
BP30	65 (260)	<10 ⁻⁶	0/5	—	—
DMS11	40 (160)	<10 ⁻⁶	0/5	—	—
BPDE7	44 (175)	<10 ⁻⁶	0/10	—	—
BP9	45 (180)	2.4×10^{-4}	5/6	88	2/5
BP3	49 (195)	4.3×10^{-4}	6/10	111	2/6
BP40	33 (130)	>1%	7/8	55	3/7
BP6	42 (170)	>1%	10/10	69	6/10
BPDE17	45 (180)	>1%	7/7	50	0/7

Immortal hamster embryo cell lines were tested for their ability to form colonies in soft agar medium at the passage number shown. 10^5 cells of each line were seeded in 3 ml DMEM plus 0.33% Noble agar (Difco) and 1 mg ml⁻¹ Bacto-Peptone in each of 10 5-cm dishes containing a 0.6% solidified agar base; these conditions did not permit any proliferation of normal diploid cells (see ref. 29). Colonies of more than ~100 cells were scored after 10–12 days at 37°C and representative examples isolated for subsequent confirmation of the heritability of the Aga⁺ phenotype. The tumorigenicity of lines was established by injecting 1-day-old syngeneic hamsters (5–10 animals per cell line) intradermally with 10^6 cells in 0.1 ml phosphate-buffered saline (PBS); animals were checked three times weekly for signs of tumour growth. When tumours at the site of injection had reached ~2 cm in diameter the animals were killed; both the primary tumours and organs suspected of carrying secondary growths were removed for histology and cell culture. Metastases were evident in a proportion of tumour-bearing animals; these involved mainly the lymph nodes, liver, pancreas and lungs. Invasion and destruction of underlying muscle was common, and occasionally extended to the diaphragm, ovary and kidney; angiogenesis associated with the primary tumour was often marked. Animals were observed for at least 300 days before being scored as negative for tumour growth. Preliminary analysis of the karyotypes of the various Sen⁻ transformants at the time of assay revealed no consistent differences between tumorigenic and non-tumorigenic lines with respect to chromosome numbers. Most lines had diploid modes plus a significant proportion of hyperdiploid cells with 1–4 extra chromosomes. A detailed study of G-banded karyotypes is now being done.

Table 3 Properties of Aga⁺ mammalian cell variants induced by carcinogens independently of immortality

Treatment	Frequency of Aga ⁺ clones per seeded cell	No. of Aga ⁺ clones isolated	No. of clones re-plated in agar	Re-plated Aga ⁺ frequency × 10 ⁵ (±s.d.)
<i>a, Syrian hamster embryo cells</i>				
DMSO (control)	<5.0 × 10 ⁻⁷	—	—	—
MNUA (30 µg ml ⁻¹)	7.0 × 10 ⁻⁶	8	0*	—
BPDE (0.15 µg ml ⁻¹)	1.3 × 10 ⁻⁵	12	0*	—
<i>b, Human fetal lung fibroblasts</i>				
DMSO (control)	<1.0 × 10 ⁻⁷	—	—	—
BPDE (single treatment, of 0.2 µg ml ⁻¹)	<1.0 × 10 ⁻⁷	—	—	—
BPDE (0.2 µg ml ⁻¹ + 4 sequential treatments of 0.1 µg ml ⁻¹)	2.2 × 10 ⁻⁵	15	3	42 ± 8 (clone A5) 38 ± 11 (clone A7) 29 ± 8 (clone A8)

a, Secondary hamster embryo cells were treated with BPDE and MNUA as described in Table 1 legend, and were then maintained in the exponential phase of growth. Samples of cells from each culture were plated in 0.33% agar medium 3, 7 and 21 days after treatment. For both treated cultures, the maximum yield of agar colonies was obtained with cells which had been allowed 7 days for expression of the Aga⁺ phenotype; these are the values shown. Twenty large, viable colonies were isolated and transferred to 1.5-cm culture wells. *b*, Secondary normal fetal lung fibroblasts were provided by J. J. Roberts, Institute of Cancer Research. At pd 10, replicate cultures of 10⁷ cells in 175-cm² flasks were treated with 0.2 µg ml⁻¹ BPDE. After 2 weeks without subculture, but with medium changes every 3 days, the cells had recovered sufficiently to produce confluent cultures. We detected no Aga⁺ variants at this stage in a total of 10⁷ cells plated in agar. Therefore, some cultures received four additional treatments of BPDE (0.1 µg ml⁻¹) evenly spread over an 8-week period. At pd 38, rare Aga⁺ variants were detected and 15 colonies (A1–A15) were isolated; cells derived from three of these (clones A5, A7 and A8) were re-plated in agar and simultaneously tested for tumorigenicity in nude (*nu/nu*) mice. Groups of four mice were injected intradermally with 5 × 10⁶ cells in 0.1 ml PBS and, as a positive control, animals in one group each received 1 × 10⁶ Sen⁻ Aga⁺ hamster cells. Animals were observed twice a week for signs of tumours.

* Most of the cultures from Aga⁺ hamster clones possessed sufficient residual division potential to form a confluent monolayer. However, all 20 cultures senesced after two subcultures.

appearance of anchorage-independent (Aga⁺) variants in Sen⁻ cell lines (Table 2). So far we have observed no lines having a Sen⁻ Aga⁺ phenotype which are non-tumorigenic, nor have we obtained any tumorigenic lines which do not form colonies in soft agar. Moreover, cells derived from all nine fibrosarcomas that we induced in hamsters by single intradermal injections of BP (100% tumour incidence; tumours excised when ≥2 cm in diameter after 109–167 days) possessed a Sen⁻ Aga⁺ phenotype in culture without entering an obvious crisis after outgrowth from tumour explants.

Several workers^{19–21} have published evidence supporting a mutational mechanism for the induction of anchorage independence. Thus, a possible explanation for the progression of our hamster lines is that rare Aga⁺ variants arise in cultures through spontaneous mutation and that cells with the new phenotype (Sen⁻ Aga⁺) then drift through the population due to a growth advantage²¹. If this is the case, then the variable rates of appearance of Aga⁺ cells could be due to variable spontaneous mutability of the individual cell lines. To test this hypothesis, we are now examining spontaneous mutation rates at the *hprt* locus, in Aga⁺, pre-Aga⁺ and non-progressing Sen⁻ hamster dermal cell lines.

Evidence supporting the notion that immortality is an essential step in malignant transformation was obtained from experiments with secondary hamster embryo cells in which Aga⁺ variants were induced by BPDE and MNUA treatment (Table 3a). Large, viable-looking colonies were picked and transferred to 1.5-cm diameter multi-dish wells (Nunc), and 20 dividing monolayer cultures were established successfully. However, although we were able to transfer the cells to larger culture vessels at least once, all 20 isolates underwent typical senes-

cence before sufficient cells could be accumulated for repeat anchorage independence assays or tumorigenicity studies. To examine the properties of Sen⁻ Aga⁺ cells having a longer residual lifespan, we carried out similar experiments with fibroblasts derived from human fetal lung; in our hands these cells regularly underwent >100 pd in medium supplemented with selected batches of FCS (also found by Pooley *et al.*²²). We were able to induce Aga⁺ variants of these cells after multiple treatments with BPDE (Table 3b). When cells grown from isolated agar colonies were re-plated in agar, their cloning efficiency was invariably more than an order of magnitude higher than that of the original carcinogen-treated mass cultures. Despite this, these Aga⁺ cells were non-tumorigenic in athymic (*nu/nu*) mice in conditions where a Sen⁻ Aga⁺ hamster cell line formed progressively growing tumours at one-fifth of the cell inoculum; in the case of the Aga⁺ human cells we observed only small nodules which appeared after 2–3 weeks and regressed 1–2 weeks later. Moreover, the cells in culture possessed senescence characteristics similar to their Aga⁺ counterparts (see also ref. 23), ceasing proliferation after a further 12–15 pd. Therefore, in both of the above studies, it was possible to induce the Aga⁺ phenotype independently of immortality but this change alone did not endow the cells with the ability to form truly malignant tumours. Numerous attempts by ourselves and others to achieve immortal transformation of human fibroblasts by chemical carcinogens have proved unsuccessful, although two clear cases have been reported^{15,24}.

The most popular interpretation of the biological significance of the limited culture lifespan of fibroblasts is that it represents ageing at the cellular level¹, and the possibility that accumulation of errors in DNA or proteins is responsible²⁵ has received much attention. An alternative idea²⁶ is that, rather than being a manifestation of senescence, finite lifespan *in vitro* may be a process analogous to terminal differentiation. There are, as yet, insufficient data available to support conclusively either theory. Nevertheless, it is clear from our results that normal mammalian cells can be induced to escape from their commitment²⁷ to senescence by exposure to various mutagenic^{17,18} carcinogens, and that the resulting capacity for infinite multiplication precedes (and may be necessary for) malignant transformation. That such a potential for unlimited division provides the basis for the evolution of carcinogen-induced malignancy *in vivo*, by permitting repeated selection of rare variants having increasingly autonomous properties, is, we believe, an appealing hypothesis which warrants further study.

This work was supported by grants from the Cancer Research Campaign/MRC (G80/1094), the US NCI (RO1-CA-25807-04) and Gallaher Ltd. We thank Mrs M. White, Ms J. Amos and Mr T. A. Slade for technical assistance.

Received 21 June; accepted 3 August 1982.

- Hayflick, L. & Moorhead, P. *Expt Cell Res.* **25**, 585–621 (1961).
- Ponten, J. *Biochim. biophys. Acta* **458**, 397–422 (1976).
- Barrett, J. C. *et al. Cancer Res.* **37**, 3815–3823 (1977).
- Daniel, C. W., Aidells, B. D., Medina, D. & Faulkin, L. J. *Fedn Proc.* **34**, 64–67 (1975).
- Hellman, S., Botnick, L. E., Hannon, E. C. & Vignuelle, R. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 490–494 (1978).
- Klein, G. *Cancer Res.* **19**, 343–358 (1959).
- DiPaolo, J. A., Nelson, R. L. & Donovan, P. J. *J. natn. Cancer Inst.* **46**, 171–181 (1971).
- Sinkovics, J. G., Gyorkey, F., Kusyk, C. & Siciliano, M. J. *Meth. Cancer Res.* **14**, 243–323 (1978).
- Klein, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2442–2446 (1979).
- Rheinwald, J. G. & Beckett, M. A. *Cancer Res.* **41**, 1657–1663 (1981).
- Berwald, Y. & Sachs, L. *J. natn. Cancer Inst.* **35**, 641–661 (1965).
- DiPaolo, J. A., Nelson, R. L. & Donovan, P. J. *Cancer Res.* **31**, 1118–1127 (1971).
- Barrett, J. C. & Ts'o, P. O. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3761–3765 (1978).
- Franks, L. M. & Wilson, P. D. *Int. Rev. Cytol.* **48**, 55–139 (1977).
- Kakunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1334–1338 (1978).
- Lewis, L. *et al. J. Cell Sci.* **53**, 21–36 (1982).
- Newbold, R. F. & Brookes, P. *Nature* **261**, 52–54 (1976).
- Newbold, R. F., Warren, W., Medcalf, A. S. C. & Amos, J. *Nature* **283**, 596–599 (1980).
- Bouck, N. & DiMayorca, G. *Nature* **264**, 722–727 (1976).
- Bellet, A. J. D. & Younghusband, H. B. *J. cell. Physiol.* **101**, 33–48 (1979).
- Marin, G. *Expt Cell Res.* **125**, 31–36 (1980).
- Pooley, J. A., Schuman, R. F. & Pienta, R. J. *In Vitro* **14**, 405–412 (1978).
- Freedman, V. H. & Shin, S. J. *J. natn. Cancer Inst.* **58**, 1873–1875 (1977).
- Namba, M., Nishitani, K. & Kimoto, T. *Japan J. exp. Med.* **48**, 303–311 (1978).
- Orgel, L. E. *Nature* **243**, 441–445 (1973).
- Bell, E. *et al. Science* **202**, 1158–1163 (1978).
- Holliday, R., Huschtscha, L. I. & Kirkwood, T. B. L. *Science* **213**, 1505–1508 (1981).
- Sun, N. *et al. Cancer Res.* **41**, 1669–1676 (1981).
- Peel, D. M. & Stanbridge, E. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3053–3057 (1981).

Noradrenaline blocks accommodation of pyramidal cell discharge in the hippocampus

D. V. Madison* & R. A. Nicoll

Departments of Pharmacology and Physiology and *Graduate Program in Neuroscience, University of California, San Francisco, California 94143, USA

The hippocampus, as well as a variety of other brain regions, is known to receive a diffuse projection of noradrenaline (NA) containing fibres which originates in the brain stem¹⁻⁴. Although there is considerable evidence for the involvement of this system in a variety of behaviours⁵⁻⁷, the precise cellular actions of NA are poorly understood. Early studies emphasized the direct inhibitory effects of NA⁸⁻¹²; more recent experiments have shown that at several sites, NA, or stimulation of NA-containing afferents, can also facilitate excitatory synaptic responses¹³⁻¹⁸. This has led to the concept that NA increases the 'signal-to-noise' ratio of neurones¹³, acting as an 'enabling' device⁴ which allows cells to respond more briskly to conventional synaptic excitation. In the olfactory bulb, NA reduces inhibitory postsynaptic potentials by a presynaptic action¹⁹, which could contribute to enhanced excitatory synaptic responses. However, in other systems, NA has been reported to enhance excitatory responses to iontophoretically applied transmitters, and it was proposed that NA increases the sensitivity of the neurone to these excitatory transmitters¹³⁻¹⁵. We report here experiments that could explain such direct effects. We have found that NA and cyclic AMP block the Ca^{2+} -activated K^{+} conductance in hippocampal pyramidal cells and that this blockade occurs at a step subsequent to the entry of Ca^{2+} into the neurone. As a consequence, the spike frequency adaptation or accommodation which normally occurs with depolarizing stimuli is severely reduced. Thus, NA greatly increases the number of spikes elicited by a depolarizing stimulus.

Rat hippocampal slices were prepared and superfused as previously described²⁰. The normal superfusion medium consisted of 116.4 mM NaCl, 5.4 mM KCl, 2.5 mM CaCl_2 , 1.3 mM MgSO_4 , 26.2 mM NaHCO_3 , 0.92 mM Na_2HPO_4 and 11 mM glucose. The recording microelectrode, filled with 2 M KMeSO₄ (90–120 MΩ), was placed in the pyramidal cell layer under visual guidance. Pyramidal cells were activated directly by passing depolarizing current through the recording electrode using a bridge circuit. Drugs were applied by addition to the superfusing medium or, in a few experiments, by iontophoresis. Iontophoretic pipettes were filled with either 1 M sodium glutamate (pH 8) or 1 M NA (pH 4.5). The findings reported here are based on recordings from over 100 cells with resting membrane potentials of greater than -55 mV.

We have found that when CA1 pyramidal cells are excited by prolonged (for example, 600 ms) depolarizing current pulses, they respond with an initial burst of spikes, after which the cell remains silent for the duration of the pulse (Fig. 1a). Increasing the stimulus current from threshold increases the number of spikes until a maximum of six to eight spikes is reached. This spike frequency adaptation, or accommodation, is markedly attenuated by the addition of NA (1–10 μM) to the superfusate ($n = 24$). Although little change in either the passive membrane properties of the cell or in the firing threshold usually occurs, NA causes the cell to fire throughout the depolarizing current pulse (Fig. 1a). For strong depolarizing currents, pyramidal cells fired 6.8 ± 1.7 (mean $\pm$ s.d.) spikes to a 650-ms pulse in control conditions, whereas in 10 μM NA the same cells fired 16.8 ± 2.7 spikes ($n = 14$). Because these neurones fire continuously during depolarizing stimuli in the presence of NA, the number of spikes elicited is determined by the duration of the depolarizing pulse. In those few cells which were hyperpolarized by 10 μM NA (see refs 11, 12), the response to threshold depolarizing pulses was inhibited, yet the response to strong current pulses was markedly augmented. The block of accommodation by NA is also seen when pyramidal cells are depolarized by glutamate applied iontophoretically (Fig. 1c). In the presence of NA, the cells fire throughout the depolarization induced by glutamate ($n = 9$).

Fig. 1 Noradrenaline and cadmium block calcium-activated potassium responses and prolong spike discharge in hippocampal pyramidal cells. The responses in *a* to *c* are all from the same pyramidal cell. *a*, Film records of the response of cell to a depolarizing current pulse. The current trace is positioned below the voltage trace. The response in the middle column was obtained 4 min after the addition of 10 μM NA to the bathing solution. The membrane potential did not change during the application of NA. The records in 'Wash' were obtained 30 min after returning to the control solution. *b*, Simultaneously recorded chart records of the response shown in *a* (*b*₁), and response to a short (100 ms) depolarizing current pulse (*b*₂). *c*, Film records of response to 30 nA of iontophoretically applied glutamate. Membrane potential, -58 mV. The gain of *c* is the same as in *a*. *d*, Chart records from another cell showing responses to a long (*d*₁) and short (*d*₂) depolarizing current pulse recorded at different gains. The calcium antagonist cadmium (0.1 mM) was applied for 8 min. The records in 'Wash' were obtained 32 min after returning to the control bathing solution. The spikes are attenuated by the pen recorder. The time calibration is the same as in *b*. The gain for *d*₂ is the same as in *b*, while *d*₁ is half the gain of *b*. Membrane potential, -57 mV.

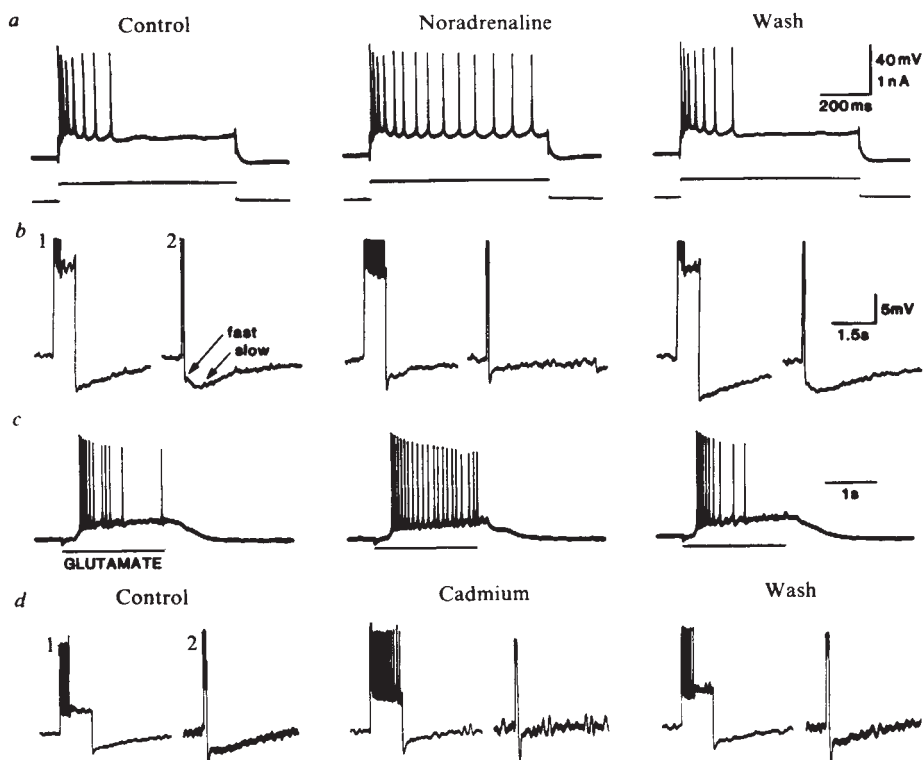
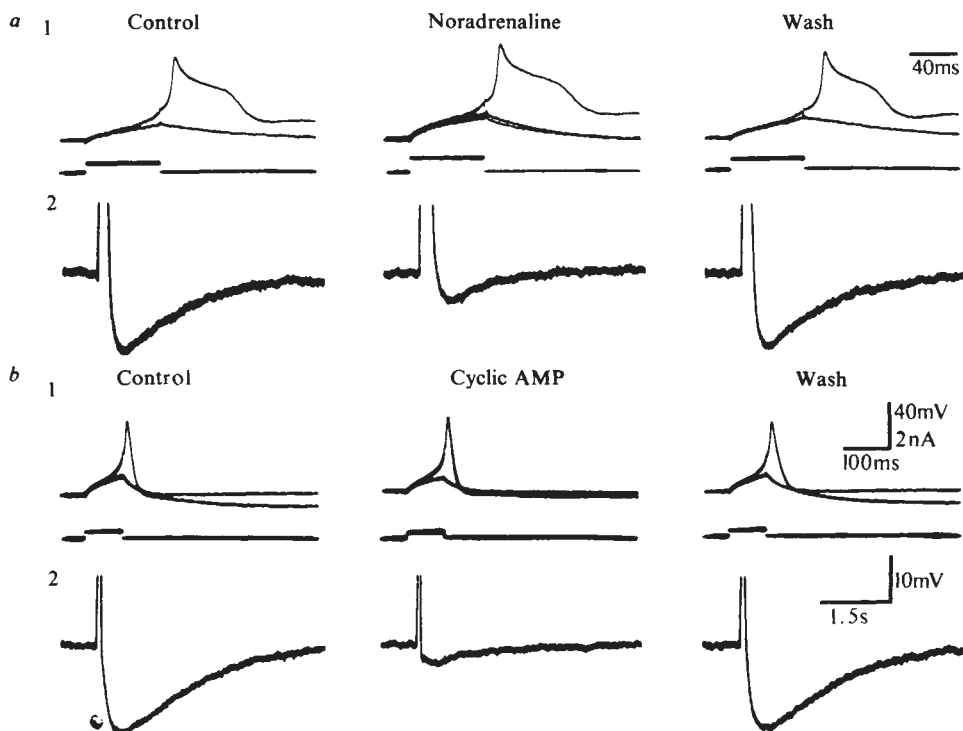


Fig. 2 Noradrenaline and cyclic AMP block calcium-activated potassium responses without reducing calcium spikes. *a*, The top traces (a_1) are film records of the Ca^{2+} spike evoked in a pyramidal cell and the corresponding current pulses. The bottom traces (a_2) are the simultaneously recorded chart records. The preparation was bathed with a solution containing $1\ \mu\text{M}$ tetrodotoxin and $5\ \text{mM}$ tetraethylammonium. The response in the middle column was obtained 7 min after starting the iontophoretic application of NA ($200\ \text{nA}$). In this cell the threshold for spike initiation was slightly increased in NA but this was not a typical finding. The response in 'Wash' was obtained 15 min after termination of NA iontophoresis. Membrane potential $-56\ \text{mV}$. *b* Shows similar responses to those in *a*, but from another cell. The record in the middle column was obtained 9 min after the addition of $1\ \text{mM}$ bromocyclic AMP to the superfusate. The record in 'Wash' was obtained 66 min after returning to the control superfusate. Membrane potential $-62\ \text{mV}$. The calibrations in *a* are the same as in *b* except for the time calibration in a_1 .



We have examined possible mechanisms underlying the ability of NA to block accommodation. A hyperpolarization of a few seconds' duration follows depolarizing current pulses (Fig. 1*b*). This hyperpolarization is due to an initial, fast, K^+ conductance followed by a slow, Ca^{2+} -activated K^+ conductance²¹⁻²³. These two components are most clearly seen in responses to short current pulses (Fig. 1*b*₂). NA reduces the late hyperpolarization but not the initial fast hyperpolarization ($n = 35$). Both the late, Ca^{2+} -activated K^+ conductance and the ability of the cell to accommodate, recover with the same time course following NA washout, suggesting that the failure of the cells to accommodate in the presence of NA is due to block of the Ca^{2+} -activated K^+ conductance. This association is strengthened by experiments in which the Ca^{2+} -activated K^+ conductance was blocked by other means. Thus, like NA, accommodation is attenuated by application of the calcium antagonist cadmium ($0.1\ \text{mM}$, added to control medium; Fig. 1*d*, $n = 10$), or low calcium solution ($2.5\ \text{mM}$ MgCl_2 substituted for CaCl_2 in control medium, $n = 3$), or recording from the cell with an electrode filled with EGTA ($0.2\ \text{M}$ EGTA plus $2\ \text{M}$ KMeSO_4 , $n = 4$). For the experiment using cadmium, in control conditions the cells fired 6.0 ± 2.2 spikes during a large 600-ms pulse, whereas in cadmium they fired 19.0 ± 4.4 spikes ($n = 10$). Furthermore, NA added to cadmium-treated cells had no additional effect.

In sympathetic ganglia, NA blocks the Ca^{2+} -activated K^+ conductance by blocking Ca^{2+} entry into the cell. This effect is mediated by an α -adrenergic receptor²⁴. NA also reduces Ca^{2+} currents in embryonic sensory neurones²⁵. To see if a similar mechanism might explain our results with NA in pyramidal cells, we have examined the effects of NA on Ca^{2+} spikes generated in the presence of tetrodotoxin ($1\ \mu\text{M}$) and tetraethylammonium ($5\ \text{mM}$). NA ($10\ \mu\text{M}$) did not reduce the Ca^{2+} spike, although the subsequent hyperpolarization was severely depressed (Fig. 2*a*). The average reduction of the after-hyperpolarization by $10\ \mu\text{M}$ NA was 87.4 ± 18.8 ($n = 14$). In fact, in the presence of NA, the calcium spike in many cells was prolonged (see Fig. 2*a*), and the cells often fired repetitive Ca^{2+} spikes. These results strongly suggest that NA blocks the Ca^{2+} -activated K^+ conductance at some step that occurs after Ca^{2+} entry into the cell.

The effect of NA on Ca^{2+} -activated K^+ conductance in pyramidal cells is not affected by the α -receptor antagonist

phentolamine ($20\ \mu\text{M}$, $n = 3$), but is antagonized by 10 – $20\ \mu\text{M}$ propranolol, a β -antagonist ($n = 7$). In addition, the β -agonist isoprenaline is approximately 10 times more potent than NA in reducing the Ca^{2+} -activated K^+ conductance ($n = 7$), while the α -agonist phenylephrine is ineffective at $10\ \mu\text{M}$ ($n = 4$). As most β -receptors exert their effects via cyclic AMP, we tested the cyclic AMP analogue, 8-bromo-cyclic AMP, 0.1 – $1\ \text{mM}$, and found that its effects mimic exactly those of NA. Thus, it strongly antagonizes the Ca^{2+} -activated K^+ response both in normal medium ($n = 14$) and in medium containing tetrodotoxin and tetraethylammonium (Fig. 2*b*, $n = 4$). This occurs without any reduction of Ca^{2+} spikes (Fig. 2*b*). At a concentration of $1\ \text{mM}$, the average reduction in the Ca^{2+} -activated K^+ response exceeded 80% ($n = 15$). It has recently been reported that 8-bromo-cyclic AMP augments the Ca^{2+} -activated K^+ response in guinea pig hippocampal pyramidal cells²⁶. We have only seen large reductions in this potential and, except for species difference, are unable to explain this discrepancy. 8-Bromo-cyclic AMP also markedly reduces accommodation of cell firing ($n = 5$, not illustrated). In control conditions, the cells fired 7.0 ± 2.0 spikes to strong 600-ms depolarizing pulses, whereas in the presence of $1\ \text{mM}$ of the analogue, the cells fired 16.7 ± 6.7 spikes ($n = 3$). The effect of the cyclic AMP analogue is not mimicked by adenosine ($10\ \mu\text{M}$, $n = 5$), indicating that the effects are not mediated by adenosine receptors. Recording electrodes were also filled with cyclic AMP ($100\ \text{mM}$ in $2\ \text{M}$ KMeSO_4) for intracellular injections. The number of spikes elicited from cells recorded with such electrodes (16.4 ± 2.8 spikes, $n = 5$) is similar to that observed after the extracellular application of the 8-bromo analogue. Forskolin, a diterpene compound which specifically and directly stimulates adenylate cyclase activity in a number of systems²⁷, also mimics the action of NA. In the presence of forskolin ($10\ \mu\text{M}$), a 600-ms depolarizing pulse elicited 18.4 ± 4.3 spikes ($n = 6$). The precise site at which cyclic AMP acts to reduce the Ca^{2+} -activated K^+ conductance is unclear, but it presumably involves some step subsequent to Ca^{2+} entry into the cell. One possible mechanism might be a cyclic AMP-induced alteration in a Ca^{2+} binding protein.

Cyclic AMP or the catalytic subunit of cyclic AMP-dependent protein kinase have been shown to modulate both voltage-sensitive and calcium-activated potassium conductances in a number of invertebrate neurones²⁸⁻³². Our results along with

these others emphasize the importance of both voltage-sensitive and calcium-activated potassium channels as targets for cyclic AMP and transmitter modulation of neuronal excitability.

Thus, our findings show that NA and cyclic AMP profoundly augment the pyramidal cell response to depolarization produced either by current injection or by iontophoresis of glutamate. NA seems to exert this effect by selectively blocking a Ca^{2+} -activated K^{+} conductance. This augmentation can be demonstrated even when NA hyperpolarizes the membrane and depresses the response to threshold current pulses. Such a combined action of suppressing weak inputs and enhancing strong inputs is consistent with the concept advanced by Woodward *et al.*¹³, that NA increases the signal-to-noise ratio in central nervous system neurones. Hippocampal neurones treated with NA thus bear a marked resemblance to the 'quiet readiness' of neurones during arousal³³, a condition associated with an increased firing rate of NA-containing locus coeruleus neurones^{6,7}.

We thank Dr H. Bourne for his interest in this work and for his many helpful suggestions, also Drs Z. Hall and M. P. Stryker for their comments on the manuscript. This work was supported by NIH predoctoral fellowship GM-07449 to D.V.M. and grant NS 15764; RCDA NS 00287 and the Klingenstein Fund to R.A.N.

Received 28 June; accepted 20 August 1982.

- Pickel, V., Segal, M. & Bloom, F. E. *J. comp. Neurol.* **155**, 15–42 (1974).
- Nakamura, S. & Iwama, K. *Brain Res.* **99**, 372–376 (1975).
- Loy, R., Koziell, D. A., Lindsey, J. D. & Moore, R. Y. *J. comp. Neurol.* **189**, 699–710 (1980).
- Moore, R. Y. & Bloom, F. E. *Rev. Neurosci.* **2**, 113–168 (1979).
- Mason, S. T. *Prog. Neurobiol.* **16**, 263–303 (1981).
- Aston-Jones, G. & Bloom, F. E. *Neuroscience* **1**, 876–886 (1981).
- Aston-Jones, G. & Bloom, F. E. *Neuroscience* **1**, 887–900 (1981).
- Bloom, F. E. in *Psychopharmacology—A 20 Year Progress Report* (eds Lipton, M. E., Killam, K. C. & Di Mascio, A.) 131–142 (Raven, New York, 1978).
- Hoffer, B. J., Siggins, G. R. & Bloom, F. E. *Brain Res.* **25**, 523–534 (1971).
- Segal, M. & Bloom, F. E. *Brain Res.* **107**, 513–525 (1976).
- Segal, M. *Brain Res.* **206**, 107–128 (1981).
- Langmoen, I. A., Segal, M. & Andersen, P. *Brain Res.* **208**, 349–362 (1981).
- Woodward, D. J., Moises, H. C., Waterhouse, B. D., Hoffer, B. J. & Freedman, R. *Fedn Proc.* **38**, 2109–2116 (1979).
- Rogawski, M. A. & Aghajanian, G. K. *Nature* **287**, 731–734 (1980).
- Waterhouse, B. D., Moises, H. C. & Woodward, D. J. *Neuropharmacology* **20**, 907–920 (1981).
- Kayama, Y., Negi, T., Sugitani, M. & Iwama K. *Neuroscience* **7**, 655–666 (1982).
- Mueller, A. L., Hoffer, B. J. & Dunwiddie, T. V. *Brain Res.* **214**, 113–126 (1981).
- Moises, H. C., Waterhouse, B. D. & Woodward, D. J. *Brain Res.* **222**, 43–64 (1981).
- Jahr, C. E. & Nicoll, R. A. *Nature* **297**, 227–229 (1982).
- Nicoll, R. A. & Alger, B. E. *J. Neurosci. Meth.* **4**, 153–156 (1981).
- Hotson, J. R. & Prince, D. A. *J. Neurophysiol.* **43**, 409–419 (1980).
- Alger, B. E. & Nicoll, R. A. *Science* **210**, 1122–1124 (1980).
- Schwartzkroin, P. A. & Stafstrom, C. E. *Science* **210**, 1125–1126 (1980).
- Horn, J. P. & McAfee, D. A. *J. Physiol., Lond.* **301**, 191–204 (1980).
- Dunlap, K. & Fischbach, G. D. *J. Physiol., Lond.* **317**, 519–535 (1981).
- Benardo, L. S. & Prince, D. A. *J. Neurosci.* **2**, 415–423 (1982).
- Seamon, K. & Daly, J. W. *J. biol. Chem.* **256**, 9799–9801 (1981).
- Kaczmarek, L. R. & Strumwasser, F. *Neurosci. Abstr.* **7**, 932 (1981).
- DePeyer, J. E., Cachelin, A. B., Levitan, I. B. & Reuter, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4207–4211 (1982).
- Livingstone, M. S. & Hubel, D. H. *Nature* **291**, 554–561 (1981).

Retroviruses induce granulocyte-macrophage colony stimulating activity in fibroblasts

Mark J. Koury & Ian B. Pragnell

Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD, UK

Murine retroviruses which cause sarcomas also induce abnormal proliferation of haematopoietic cells leading to death in newborn^{1–5} and, in some cases, adult⁴ mice. A murine sarcoma virus is a complex of a replication-defective sarcoma virus and a replication-competent murine leukaemia virus (MuLV). Cloned MuLV isolates also cause abnormal haematopoietic cell proliferations^{6–9}. It is unknown whether such proliferative

responses are direct effects of retroviruses on haematopoietic progenitor cells or indirect effects on other cells, which in turn provide the stimuli for the proliferation of haematopoietic cells. We demonstrate here that specific fibroblastic cell lines produce low levels of granulocyte-macrophage colony stimulating activity (GM-CSA) and that infection of certain fibroblast lines with murine retroviruses results in large increases in GM-CSA. Other cell lines which do not produce detectable CSA when uninfected do not produce it after retroviral infection.

Granulocyte-macrophage colony forming cells (GM-CFC) are haematopoietic progenitor cells which respond to GM-CSA by proliferation and differentiation *in vitro*¹⁰ yielding either pure or mixed colonies of granulocytes and macrophages. Sera of mice having virus-induced leukaemias contain large amounts of GM-CSA compared with sera of uninfected controls^{11,12}. Adult mice injected with myeloproliferative virus (MPSV), one of the sarcoma viruses, develop acute splenomegaly, anaemia and increased numbers of blood granulocytes and granulocytic precursors^{13,14}. These haematological changes are accompanied by fibroblastic proliferation primarily in the bone marrow and

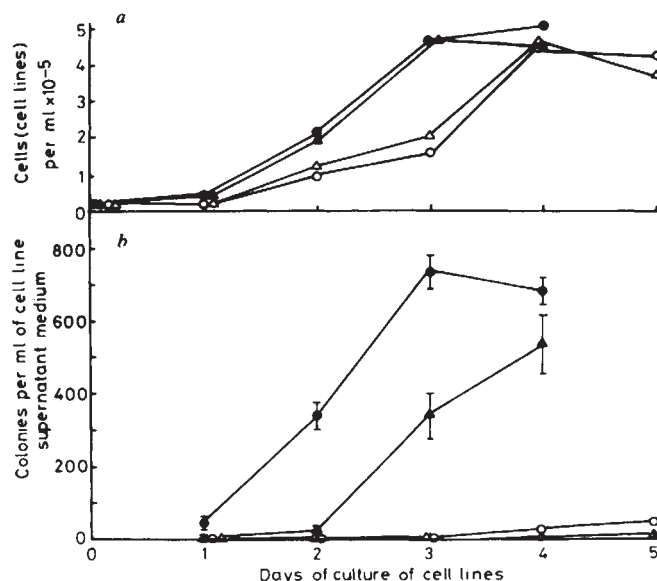


Fig. 1 Effect of sarcoma virus infection on growth and GM-CSA in supernatant medium of cell line cultures. Normal rat kidney (NRK)¹⁵ and mouse embryo (MMCE)¹⁶ cell lines were infected with myeloproliferative virus (MPSV). Infection was confirmed by reverse transcriptase¹⁷ and fibroblast-transforming¹⁸ activities of supernatant culture media. Infected cells and their uninfected controls were cultured in Dulbecco's modified Eagle's minimal essential medium (DMEM) plus 10% fetal calf serum (FCS). Replicate 25-cm² flasks of 1.5×10^5 cells in 7 ml were incubated at 37 °C in 5% CO₂ in air. At daily intervals the medium plus any nonadherent cells of a replicate culture were collected. After centrifugation to pellet any cells, the supernatant medium was filtered through 0.2- μm^2 filters and stored at 4 °C. Adherent cells were collected after treatment with trypsin, added to the centrifuged nonadherent cells and total cells were counted. Cell concentrations per ml are shown in a. The supernatants were assayed for the presence of GM-CSA with bone marrow cells from 8–12-week-old female BALB/c mice. The assays were done by culturing, in 16-mm diameter wells, 0.5 ml aliquots of 5×10^4 nucleated marrow cells per ml in supplemented DMEM¹⁹. The medium also contained 0.8% methylcellulose, 20% FCS, 100 U ml⁻¹ penicillin, 100 μg ml⁻¹ streptomycin and 20% cell line supernatant. Colonies of ≥ 50 cells were scored after 7 days incubation at 37 °C in humidified air plus 5% CO₂. Colonies were identical to the types found with spleen cell-conditioned medium as a source of GM-CSA¹⁰. The granulocytes and macrophages were identified by their morphology and by staining with Wright's or Leishman's stain in clotted, fixed cultures²⁰. In cytocentrifuge preparations of washed cultured cells, the morphological studies were confirmed by the staining of the macrophages with α -naphthylacetate esterase and the granulocytes with chloroacetate esterase. CSA per ml of supernatant medium was determined from quadruplicate cultures, (b). ●, MPSV-infected NRK; ▲, uninfected NRK; ○, MPSV-infected MMCE; △, uninfected MMCE. All results are ± 1 s.d.

Table 1 Effects of sarcoma virus and MuLV on production of GM-CSA

Cell line/infection	Supernatant activities		
	GM-CFC stimulation (colonies per ml)	Reverse transcriptase	Fibroblast focus formation
NRK			
Uninfected	50 ± 10	—	—
MPSV complex	540 ± 30	+	+
MPSV complex (56 °C)	590 ± 50	—	—
MPSV (NP)	400 ± 50	—	—
ts MPSV (NP) (39 °C)	480 ± 70	—	—
ts MPSV (NP) (33 °C)	550 ± 80	—	—
MuLV-F	280 ± 50	+	—
Moloney complex	390 ± 70	+	+
Moloney (NP)	300 ± 60	—	—
MuLV-Mol	190 ± 40	+	—
Kirsten (NP)	340 ± 30	—	—
SC-1			
Uninfected	110 ± 40	—	—
MPSV complex	590 ± 50	+	+
MPSV complex (56 °C)	540 ± 70	—	—
MuLV-F	470 ± 80	+	—
NIH/3T3			
Uninfected	50 ± 40	—	—
Harvey (NP)	270 ± 40	—	—
FRE			
Uninfected	0	—	—
MPSV complex	0	+	+
MuLV-F	0	+	—
MMCE			
Uninfected	0	—	—
MPSV complex	0	+	+
MuLV-F	0	+	—
Mixed supernatants			
10% FRE (MPSV) + 10% SC-1 (MPSV)	420 ± 50		
10% MMCE (MPSV) + 10% SC-1 (MPSV)	480 ± 30		
10% Culture medium + 10% SC-1 (MPSV)	440 ± 60		
20% Culture medium	0		

Continuous cell lines were infected with sarcoma virus complex, infected with MuLV-F (clone 643/22N; ref. 21), or were isolated containing a defective sarcoma virus component [that is, non-producer cells (NP)]. In the Moloney sarcoma complex the helper virus is Moloney MuLV (MuLV-Mol; ref. 21). Some non-producer lines were superinfected with MuLV-F. Each cell in the cell line cultures was infected as most cell lines were either clones from single infected cells or were chronically infected for at least 2 months before the experiment. The SC-1 and NIH 3T3 cell lines have been described elsewhere^{22,23}. The Fischer rat embryo (FRE) line was a gift from E. Scolnick. The non-producer lines of MPSV, Moloney, Kirsten, and Harvey viruses have been described²¹, as has the temperature sensitive mutant (ts 159/5/1)²⁴. The cells were plated in flasks and cultured as described in Fig. 1 legend. After 2 days of culture when the cells were subconfluent, the medium was changed and they were allowed to grow to confluence for the following 2 days. After the second 2-day period, the supernatant medium was collected, processed, and GM-CSA determined from quadruplicate cultures as described in Fig. 1 legend. Supernatants were collected from cultures having similar numbers of cells. Adjustments for cell numbers were made in the supernatant assays of ts mutants due to altered growth rates at the permissive (33 °C) and nonpermissive (39 °C) temperatures for sarcoma phenotypic expression. Heat-treated samples were raised to 56 °C for 30 min. Reverse transcriptase and focus-forming assays were done on supernatant media (see Fig. 1 legend). All results are ± 1 s.d.

later in the spleen¹³. Sera of mice during these acute stages of MPSV infection show large increases in GM-CSA compared with controls (our unpublished results). However, *in vivo* studies do not reveal the mechanism by which retrovirus infection results in these pathological changes. Because of its effects on fibroblast and granulocyte proliferation and on serum GM-CSA, we initially used MPSV to investigate the relationship between infection of fibroblasts by sarcoma virus and the induction of haematopoietic proliferation.

Cell lines infected with MPSV were compared with uninfected controls for growth and for GM-CSA production (Fig. 1). Similar growth was observed in infected and uninfected cultures (Fig. 1a). During the first two days of culture, the supernatant medium of normal rat kidney (NRK) cells contained slight amounts of GM-CSA, while infected NRK cells had much greater amounts of GM-CSA in their supernatant medium (Fig.

1b). Supernatants of infected and uninfected NRK cells cultured for 4 days stimulated the growth of similar numbers of colonies (Fig. 1b), but the colonies stimulated by infected cell supernatants contained several hundred of cells per colony while those stimulated by uninfected cell supernatants contained 50–100 cells. This difference in colony size suggested a greater amount of GM-CSA in the infected cell supernatants. The increased amount of GM-CSA in the 4 day-infected cell supernatants was confirmed by dose titration as shown in Fig. 2. In contrast to NRK cells, a mouse embryo cell line, MMCE, had negligible GM-CSA in the supernatant medium at all times, whether the cells were infected or uninfected (Fig. 1b). This lack of GM-CSA occurred despite MMCE cell growth to concentrations sufficient for large amounts of GM-CSA production by NRK cells.

Other cell lines and viruses were also examined (Table 1). In this study we found that some cell lines had constitutive production of GM-CSA while other lines did not. This ability to produce GM-CSA effective for mouse GM-CFC development was found in fibroblast lines of both adult and fetal origin and in lines of both rat and mouse origin. However, if a cell produced GM-CSA when uninfected, then infection with MPSV, MuLV, or the defective sarcoma virus (that is, non-producer state) resulted in large increases in GM-CSA. A consistent pattern in multiple experiments was that the largest increase in GM-CSA occurred with sarcoma complex infections, with slightly less of an increase for sarcoma non-producers, and the smallest increase with MuLV infections. Moloney, Kirsten, and Harvey sarcoma viruses increased GM-CSA as well as MPSV (Table 1). Conversely, if GM-CSA in the supernatant of an uninfected cell line was not detected, then infection of the line did not induce any GM-CSA after 2 days (Table 1) and even 5 days of culture (data not shown). The lack of GM-CSA in MMCE and FRE supernatants was not due to an inhibitor of colony formation, as shown by mixing experiments (Table 1).

Retroviruses did not directly stimulate colony growth as non-producer cell supernatants (which lack virus particles) and heat-treated supernatants (which have inactivated virus) both showed increased GM-CSA (Table 1). Furthermore,

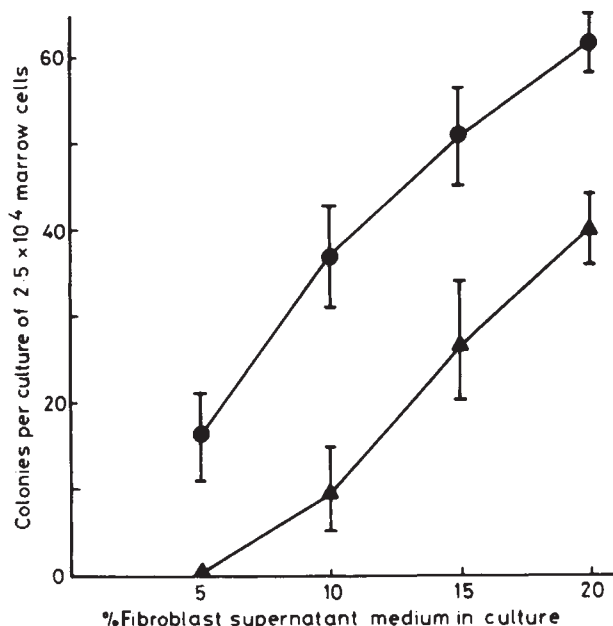


Fig. 2 Titration of GM-CSA in fibroblast line supernatants. The same day 4 supernatants of MPSV-infected and uninfected NRK cells as in Fig. 1 were assayed at various dilutions for GM-CSA. The assay was performed as described in Fig. 1 legend. All data are ± 1 s.d. for quadruplicate cultures. ●, MPSV-infected NRK; ▲, uninfected NRK.

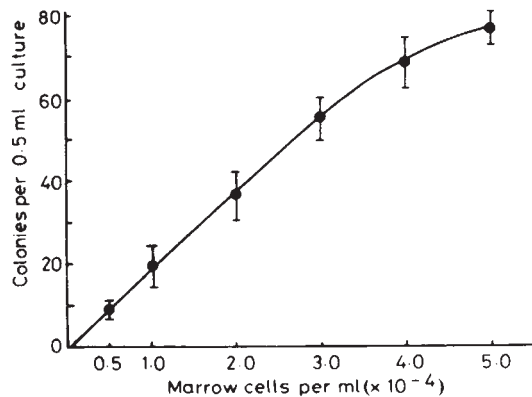


Fig. 3 Effect of decreasing the number of marrow cells on the assay of GM-CSA from infected cells. Supernatant from 2-day cultures of MPSV-infected SC-1 cells comprised 20% of the GM-CSA assay medium as described in Fig. 1 legend. Marrow cell concentrations in the assay were varied as shown. All data are ± 1 s.d. of quadruplicate cultures.

phenotypic expression of the sarcoma genome (that is, morphological transformation) was not necessary for increased GM-CSA production because cells infected with a temperature-sensitive mutant of MPSV had similar GM-CSA production at both permissive and non-permissive temperatures (Table 1). The direct GM-CSA action on GM-CFC is suggested by the linear slope intersecting zero as cell number is decreased in the assay (Fig. 3). Colonies grown with infected NRK or SC-1 cell supernatants were generally 50% macrophage, 25% granulocyte and 25% mixed.

Although it is not known how retroviruses cause the abnormal haematopoietic proliferation found in infected mice, our results suggest that large increases in the production of haematopoietic growth factors by infected fibroblasts may well play a part in the process. This indirect action of the virus on the haematopoietic progenitor cells may be further amplified in the case of the sarcoma viruses. In these infections, sarcomatous growth of fibroblasts or other mesodermal cells could result in an increased number of cells, each producing an increased amount of haematopoietic growth factor.

The studies were funded by grants from the Leukaemia Research Fund, the MRC and the Cancer Research Campaign. We thank Dr W. Ostertag for his advice about and supply of cloned cell lines and viruses and Dr T. M. Dexter for his advice about colony assays.

Differential expression of cellular oncogenes during pre- and postnatal development of the mouse

Rolf Müller*, Dennis J. Slamon†, Joanne M. Tremblay*, Martin J. Cline† & Inder M. Verma*

* Tumor Virology Laboratory, The Salk Institute, PO Box 85800, San Diego, California 92138, USA

† Division of Hematology/Oncology, Department of Medicine, UCLA Medical Center, Los Angeles, California 92004, USA

The acutely oncogenic retroviruses induce a wide spectrum of malignancies including sarcomas, carcinomas, lymphomas and leukaemias¹. They all contain sequences which are required for neoplastic transformation as an integral part of their genomes¹⁻³. These sequences, termed viral oncogenes (*v-onc*), apparently originated from the normal vertebrate genome. To date, cellular homologues of over a dozen different oncogenes of acutely oncogenic viruses have been identified¹⁻⁴. Although transcription of some *c-onc* genes has been detected in normal vertebrate cells^{1,5-7}, relatively little is known about their role in normal cell metabolism. We describe here stage- and tissue-specific expression of cellular genes homologous to the oncogenes of FBJ murine osteosarcoma virus, Abelson murine leukaemia virus and Harvey sarcoma virus during mouse prenatal and early postnatal development. Our results suggest participation of cellular oncogenes in normal developmental processes.

The FBJ murine osteosarcoma virus (FBJ-MSV) complex was isolated from a spontaneous tumour in a 260-day-old CF1 mouse⁸. Serial passage of cell-free tumour extracts in newborn mice resulted in a virus stock which induced osteosarcomas with a latency as short as 3 weeks⁹. A 55,000 molecular weight (55K) protein encoded by the FBJ-MuSV transforming gene (*v-fos*) has been described previously¹⁰. The cellular homologue of *v-fos* could be identified in DNA of normal vertebrate cells¹¹.

Relatively high levels of *c-fos*-related sequences were detected in poly(A⁺) RNA prepared from day 6, 7, 8 and 9 conceptuses containing the embryo proper and extraembryonal tissues (Figs 1a, 2a). More than 10-fold lower *fos* expression was observed in embryos of later developmental stages dissected free of extraembryonal tissues (Fig. 1a), as measured by density scanning of the autoradiogram (Fig. 1c). To determine if this dramatic change is due to a stage-specific *c-fos* expression in the embryo proper or rather a consequence of a cell-type-specific expression in an extraembryonal tissue, we analysed poly(A⁺) RNA prepared separately from day 10 embryos, extraembryonal membranes and placenta. Figure 2a shows that *c-fos*-related transcripts were found at considerably higher quantities in the placenta (PL) than in the embryo proper (E) and the extraembryonal membranes (M). Consistent with these observations is the high level of *c-fos* expression in day 18 placenta and the absence of significant amounts of *c-fos*-related transcripts in day 12 extraembryonal membranes. These findings suggest that *c-fos* expression at early developmental stages (days 6-9) probably occurs in cells which give rise to the placenta at later developmental stages. Expression of *c-fos* in the fetus increases about fivefold at later stages of development (after day 16 of gestation, Fig. 1a, c). In an attempt to identify cells expressing *c-fos* transcripts, we analysed poly(A⁺) RNA from various postnatal tissues. Figure 2b shows that low levels of *c-fos* expression were observed in all tissues investigated. However, RNA isolated from 'bone' (sternum, ribs and vertebra of 1-3 day old mice, including bone marrow, muscles and part of the pleura) and 'skin' (including the subcutaneous connective

Received 5 July; accepted 9 August 1982.

- Harvey, J. J. *Nature* **204**, 1104-1105 (1964).
- Perk, K. & Moloney, J. B. *J. natn. Cancer Inst.* **37**, 581-599 (1966).
- Kirsten, W. H. & Mayer, L. A. *J. natn. Cancer Inst.* **39**, 311-335 (1967).
- Chirigos, M. A., Scott, D., Turner, W. & Perk, K. *Int. J. Cancer* **3**, 223-237 (1968).
- Scher, C. D., Scolnick, E. M. & Siegler, R. *Nature* **256**, 225-226 (1975).
- Fischinger, P. J., Moore, C. O. & O'Connor, T. E. *J. natn. Cancer Inst.* **42**, 605-622 (1969).
- Troxler, D. H., Ruscetti, S. K., Linemeyer, D. L. & Scolnick, E. M. *Virology* **102**, 28-45 (1980).
- MacDonald, M. E., Mak, T. N. & Bernstein, A. *J. exp. Med.* **151**, 1493-1503 (1980).
- McGarry, M. P., Steeves, R. A., Eckner, R. J., Mirand, E. A. & Trudel, P. J. *Int. J. Cancer* **13**, 867-878 (1974).
- Burgess, A. W. & Metcalf, D. *Blood* **56**, 947-958 (1980).
- Robinson, W., Metcalf, D. & Bradley, T. R. *J. cell. Physiol.* **69**, 83-91 (1976).
- Metcalf, D. & Foster, R. J. *J. natn. Cancer Inst.* **39**, 1235-1245 (1967).
- LeBousse-Kerdiles, M. C., Smadja-Joffe, F., Klein, B., Caillou, B. & Jasmin, C. *Eur. J. Cancer* **16**, 43-51 (1980).
- Klein, B. et al. *J. natn. Cancer Inst.* **66**, 935-940 (1981).
- Huu, D., Rosenblum, E. N. & Zeigel, R. F. *J. Bact.* **92**, 1133-1140 (1966).
- Rapp, U. R., Keski-Oja, J. & Heine, U. *Cancer Res.* **39**, 4111-4118 (1979).
- Pragnell, I. B., Ostertag, W. & Paul, J. *Expl Cell Res.* **108**, 268-276 (1977).
- Billello, J. A., Strand, M. & August, J. T. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3234-3238 (1974).
- Kost, T. A., Koury, M. J. & Krantz, S. B. *Virology* **108**, 309-317 (1981).
- McMahon, Y. & Hankins, W. D. *Expl Hemat.* **8**, 1081-1085 (1980).
- Ostertag, W. et al. *J. Virol.* **33**, 573-582 (1980).
- Hartley, J. W. & Rowe, W. P. *Virology* **65**, 128-134 (1975).
- Todaro, G. J. & Green, H. *J. Cell Biol.* **17**, 299-313 (1963).
- Ostertag, W. et al. in *Expression of Differentiated Functions in Cancer Cells* (ed. Revoltella, R.) 401-416 (Raven, New York, 1981).

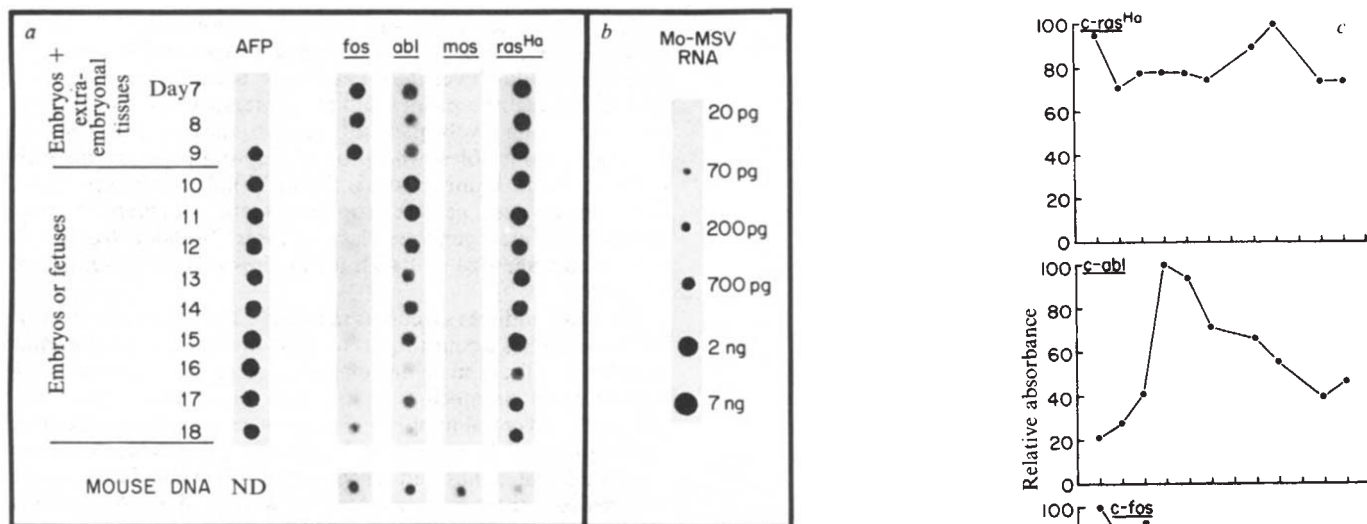


Fig. 1 *a*, Expression of cellular oncogenes during mouse prenatal development. RNA was isolated from Swiss-Webster mouse embryos/fetuses at daily intervals using the guanidine thiocyanate method²⁷. Beginning at day 10 of prenatal development (day of coital plug was taken as day 0 of prenatal development) the embryo proper was separated from the extra-embryonal membranes and placenta. Separation was not performed before this stage because of the very small size of the embryo proper. Therefore, day 6–9 embryos represent the entire conceptus as dissected from the uterine wall. RNA was selected for poly(A⁺) RNA by one cycle of chromatography on oligo(dT)–cellulose columns²⁸. Poly(A⁺) RNA was dissolved in H₂O, boiled, quick-cooled on ice, and 3 μ g (1.5 μ l) applied to sheets of nitrocellulose paper which had previously been equilibrated with 20 $\times$ SSC (1 $\times$ SSC is 0.15 M NaCl, 0.015 M sodium citrate) and air dried. After baking overnight at 80 $^{\circ}$ C the blots were prehybridized for at least 4 h at 45 $^{\circ}$ C in a buffer containing 0.75 M NaCl, 0.05 M sodium phosphate (pH 7.5), 0.005 M EDTA, 0.2% SDS, 10 mg per ml glycine, 5 $\times$ Denhard's reagent (1 $\times$ Denhard's reagent is 0.02% each of Ficoll, bovine serum albumin and polyvinylpyrrolidone), 0.25 mg of denatured herring DNA per ml and 50% formamide. Subsequently, the blots were hybridized for $\sim$ 20 h at 45 $^{\circ}$ C with 1 $\times$ 10⁶ c.p.m. of nick-translated probe per ml of hybridization buffer (same composition as prehybridization buffer except that the concentration of Denhard's reagent was decreased to 1 $\times$). The cloned oncogene fragments purified from vector sequences by preparative agarose gel electrophoresis were nick-translated²⁹ in the presence of either [³⁵S]-dATP (500 Ci mmol⁻¹, used for the α -fetoprotein (AFP)-specific probe) or [³²P]-dCTP (3,200 Ci mmol⁻¹, used for all the other probes) to specific radioactivities of $\sim$ 2 $\times$ 10⁸ c.p.m. per μ g of DNA for ³⁵S-labelled probes and 1–2 $\times$ 10⁹ c.p.m. per μ g of DNA for ³²P-labelled probes, respectively. The following fragments from cloned retroviral genomes were used as *onc*-specific probes: *fos*, BglII–PvuII (1.3 kbp) or PstI–PvuII (1.0 kbp; ref. 11); *mos*, XbaI–HindIII (1.1 kbp; ref. 30); *abl*, SmaI–BglII (~1.2 kbp; ref. 31); *ras*^{Ha}, BglII–SalI (0.46 kbp; ref. 32). After hybridization the blots were washed three times in 1 $\times$ SSC at 50 $^{\circ}$ C for about 2 h and exposed to X-ray films with intensifying screens at –70 $^{\circ}$ C for 72 h. Bottom row: hybridization of probes to total mouse genomic DNA. One microgramme of heat-denatured mouse DNA was applied to the same nitrocellulose filters used for RNA analysis and thus subjected to the same hybridization procedure, to assess the relative extent of hybridization of the nick-translated viral oncogene probes to the corresponding mouse homologue in the conditions used. ND, not done. A crucial point of the validity of the present results is the authenticity of RNA samples isolated from embryos/fetuses and extraembryonal tissues at distinct stages of prenatal development. We therefore included in all studies concerning prenatal development a mouse AFP-specific probe (pBR322–AFP2; ref. 33). AFP polypeptide is known to be expressed exclusively in visceral endoderm (starting at around day 7 of prenatal development in a few cells) and at later developmental stages also in fetal liver^{34,35}. We observed absence or very low levels of AFP-related mRNA transcripts in day 6, 7 and 8 conceptuses as compared with embryos/fetuses of later developmental stages as well as relatively low AFP expression in 10-day embryos as compared with, for example, extraembryonal membranes (Figs 1a, 2a). Our findings agree closely with those of previously published studies on protein expression^{34,35}. *b*, Sensitivity of dot blot analysis. The *v-mos*-specific probe was hybridized to varying amounts of sucrose-gradient purified 60/70 Mo-MSV RNA spotted on nitrocellulose paper. The experiment was performed essentially as described above. The detection limit was $\sim$ 20 pg of Mo-RSV RNA (5.6 kb, equivalent to $\sim$ 1 RNA transcript of $\sim$ 3 kb per cell; $\sim$ 2 μ g of poly(A⁺) RNA applied to filter). Exposure time was 48 h. *c*, Quantitation of *c-osc* expression during mouse prenatal development. The dot blot autoradiogram shown in *a* was evaluated by soft laser density scanning and the relative absorbance was plotted against the stage of development. Values were normalized to 100 for the highest level of expression in the case of each *c-osc*. Day 16 was not considered because the amount of poly(A⁺) RNA applied to the filter (*a*) was considerably lower than for all other developmental stages, as shown by a shorter exposure of the AFP blot (*a*).

tissue, muscles and part of the peritoneum) from 1–3 day old mice showed a 5–20-fold greater hybridization to the *fos*-specific probe than did RNA from all other tissues (Fig. 2b). Little transcriptional activity of *c-fos* could be detected in any haematopoietic tissue (Fig. 2b), including fetal liver (data not shown). Therefore, it appears that those cells exhibiting high levels of *c-fos* expression are not located in the bone marrow. Moreover, elevated levels of *c-fos* expression were also observed in 'skin' (Fig. 2b), suggesting that the enhanced *c-fos* transcription occurs in cells of mesenchymal origin which participate in the formation of serous membranes or connective tissues. A role of the *c-fos* gene in normal bone differentiation would be of particular interest, as its viral homologue (*v-fos*) has been implicated in the induction of osteosarcomas by FBJ-MSV in newborn mice^{8–11}.

To characterize the size of the *c-fos* transcripts, poly(A⁺) RNA isolated from embryos/fetuses of various developmental stages was analysed by agarose gel electrophoresis and Northern blotting on nitrocellulose paper (Fig. 3A). A single transcript of $\sim$ 2.2 kilobases (kb) was detected in day 7 conceptuses and day 18 fetuses, consistent with the results obtained by dot blot analysis (Fig. 1a). Similarly, a 2.2-kb *c-fos*-specific transcript was observed only in day 10 placenta, not in extraembryonal

membranes (Fig. 3A). In agreement with the dot blot analysis (Fig. 2b), a 2.2-kb transcript was also found in poly(A⁺) RNA isolated from 'bones' and 'skin' of newborn animals (Fig. 3A).

Albion murine leukaemia virus (A-MuLV) induces lymphosarcomas of pre-B-cell origin¹² and T-cell lymphomas¹³ in mice. In virus-transformed cells, a 120K protein (p120^{abl}) encoded by the A-MuLV oncogene (*v-abl*) has been identified¹⁴. In normal cells of murine haematopoietic tissues, a 150K protein (NCP150), which can be immunoprecipitated by an anti-A-MuLV tumour regressor serum, has also been reported¹⁵.

Despite significant transcriptional activity of *c-abl* throughout mouse prenatal development, a differential pattern of expression was observed. RNA transcripts related to *c-abl* were found at three- to fivefold lower concentrations in poly(A⁺) RNA isolated from day 6, 7, 8 and 9 conceptuses containing the embryo proper and extraembryonal tissues than from embryos/fetuses at day 10–11 of prenatal development which are devoid of extraembryonal tissues (Figs 1a, 1c, 2a). Accordingly, *c-abl* is expressed at about threefold higher levels in the embryo proper than in day 10 extraembryonal membranes and placenta (Fig. 2a) as determined by density scanning of the autoradiogram. Based on the dot blot analyses

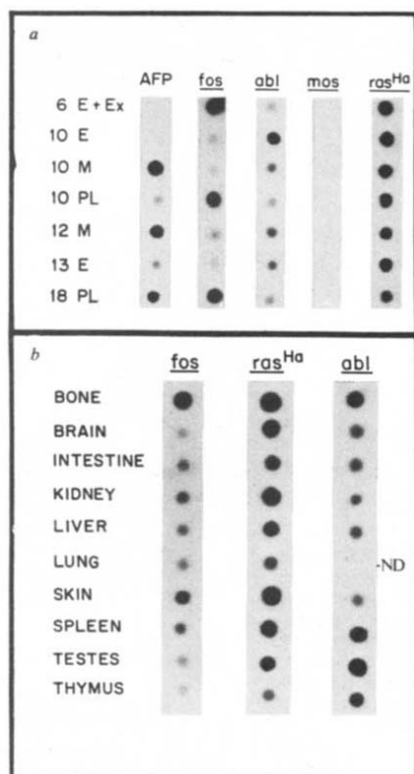


Fig. 2 *a*, Expression of *c-onc* genes in mouse embryos compared with extraembryonal tissues. The embryo proper was surgically separated from the extraembryonal membranes (that is, amnion and visceral yolk sac) and the placenta. Isolation of poly(A⁺) RNA and performance of dot blots were as described in Fig. 1*a*, legend except that 0.5 µg of twice poly(A⁺) RNA (selected twice) was spotted onto the nitrocellulose filter in each case. All probes were ³²P labelled. Blots were exposed for 1 h (AFP) and 72 h (all others), respectively. Day 6 embryos containing extraembryonal tissues (6E+Ex); day 10 embryos (10E), membranes (10M) and placenta (10PL); day 12 membranes (12M); day 13 embryo (13E); day 18 placenta (18PL), containing part of the adherent membranes. *b*, Expression of *c-onc* genes in postnatal mouse tissues. The dot blot analysis of poly(A⁺) RNA (selected once) was carried out exactly as described above. Blots were exposed for 72 h (*fos*) and 24 h (*abl*, *ras*^{Ha}), respectively. All tissues were dissected from 10-day-old mice except for bone, skin (see text) (1–3 day old mice) and testes (3-month-old mice). A potential difficulty with the quantitative evaluation of this blot is the possible variability in mRNA content of various tissues. Thus, equal amounts of poly(A⁺) RNA from different tissues have not necessarily been extracted from equal numbers of cells. Furthermore, the amount of mRNA in poly(A⁺) RNA fractions obtained from different tissues can be subject to considerable variations. Therefore, less than threefold differences on this blot autoradiogram could reflect such variations rather than differences in gene expression.

(Figs 1*a*, *c*, 2*a*), expression of *c-abl* in the fetus seems to decrease after day 11 of prenatal development. These results could suggest transcriptional activity of *c-abl* in various tissues, with at least one cell type expressing *c-abl* at significantly higher levels. This explanation is lent further support by the fact that *c-abl* is expressed in all postnatal mouse tissues examined, with spleen and thymus poly(A⁺) RNA exhibiting a stronger hybridization (~threefold) to the *abl*-specific probe than poly(A⁺) RNA from other tissues (Fig. 2*b*). By far the highest level of *c-abl* expression was observed in testes of young adult mice (Fig. 2*b*). Size analysis by agarose gel electrophoresis and Northern blotting revealed two size classes of *c-abl* transcripts of approximately 4.7 and 5.7 kb in day 6, 10 and 13 embryos as well as in postnatal thymus and spleen (Fig. 3*B*). In adult testes, an additional transcript of 3.7 kb was detected which is considerably more abundant than the 4.7- and 5.7-kb mRNAs (Fig. 3*B*). This observation may suggest that initiation or termination of *c-abl* transcription or splicing of the *c-abl*-related transcript is a precisely modulated tissue-specific mechanism.

Harvey sarcoma virus (HaSV) causes erythroleukaemia and sarcomas in rodents¹⁶. Its transforming gene (*v-ras*^{Ha}) originated from normal rat cellular sequences¹⁷. A *ras*^{Ha}-encoded protein of molecular weight 21K (p21) has been identified in

both virus-transformed and normal mouse cells^{18,19}. In the present study, *c-ras*^{Ha} was found to be expressed at considerable, but similar levels at all stages of prenatal development, both in the embryo proper and in extraembryonal tissues (Figs 1*a*, *c*, 2*a*). High levels of *c-ras*^{Ha} expression were also observed in various tissues of newborn or 10-day-old mice, particularly in 'bone', brain, kidney and 'skin' (Fig. 2*b*). Analysis of poly(A⁺) RNA by agarose gel electrophoresis and Northern blotting indicates a heterogeneous class of *c-ras*^{Ha}-related transcripts with an average size of 1.4 kb for all pre- and postnatal tissues (Fig. 3*C*).

Mo-MSV induces sarcomas in mice²⁰. The oncogene (*v-mos*) of Mo-MSV has a counterpart (*c-mos*) in the genome of normal vertebrate cells. A sustained effort to detect expression of *c-mos* sequences in uninfected mouse cells has proved unsuccessful^{21,22}. No expression of *v-mos* sequences was observed during mouse prenatal development (Figs 1*a*, 3*A*). It has recently been suggested that *c-mos* sequences are highly methylated, which may prevent their transcription²². The lack of RNA hybridization to the *v-mos*-specific probe in Fig. 1 shows that nonspecific hybridization is minimal in the dot blot analysis used in the present study.

Embryos and unfractionated tissues consist of a great variety of cell types, which may exhibit considerably different levels of *c-onc* expression. Therefore, studies pertaining to a quantitative evaluation of *c-onc* expression can only yield average values for the heterogeneous cell population of a given tissue or an embryo at a particular developmental stage. However, to obtain an estimation of the order of *c-onc* expression in mouse embryos, transcriptional activity of *c-abl* and *c-ras*^{Ha} in total embryos at day 11 of prenatal development was compared with expression of the corresponding *v-onc* genes in cell lines transformed by A-MuLV or HaSV. Figure 4 shows that the relationship between the amount of RNA applied to the filter and the signal observed on the autoradiogram is approximately linear. The levels of *v-onc* expression in the virus-transformed cell lines were found to be about 10-fold higher than expression of the corresponding *c-onc* genes in the total embryo (Fig. 4). Because one can assume that certain *c-oncs* are expressed at significantly elevated levels only in a limited number of cells, the level of expression in these particular cells is presumably much higher than the observed average value.

Significant expression of the *c-fos* gene has been detected during mouse prenatal development specifically in the placenta (Figs 1*a*, 2*a*, 3*A*). Both fetal and maternal moieties of the mouse placenta—trophoblast and decidua—are very rapidly proliferating tissues. The mouse trophoblast has been considered a 'pseudo-malignant' tissue, because it grows invasively into the uterine epithelium or any other surrounding tissue²³. It will now be of particular interest to determine if these growth properties can be correlated with the expression of a *c-onc* gene like *c-fos*. The pattern of *c-abl* expression in postnatal tissues (Figs 2*b*, 3*B*) suggests a physiological role of this *c-onc* gene in the male germ cell lineage, and possibly a different role in development of lymphatic tissues. In day 10–12 embryos, however, cell types exhibiting high levels of *c-abl* expression could not be identified, but apparently are not germ cell precursors, as the testes-specific 3.7 kb *c-abl* transcript does not seem to be present in embryos of any developmental stage (Fig. 3*B*). The role of *c-ras*^{Ha} in physiological processes remains obscure because it is expressed in embryos/fetuses of all developmental stages as well as in a variety of pre- and postnatal tissues at similar levels (Figs 1*a*, *c*, 2, 3*C*). It has been shown that *c-ras*^{Ha} linked to a retroviral long terminal repeat induces uncontrolled cell proliferation when introduced into a mouse fibroblastic cell line¹⁷. This finding suggests that a role of *c-ras*^{Ha} in normal cell proliferation is possible, but, however, may be confined to particular cell types.

In the present study, as well as in another investigation²⁴ concerned with expression of avian oncogenes, we have demonstrated differential transcriptional activity of *c-onc* genes during pre- and early postnatal development of the mouse. Our

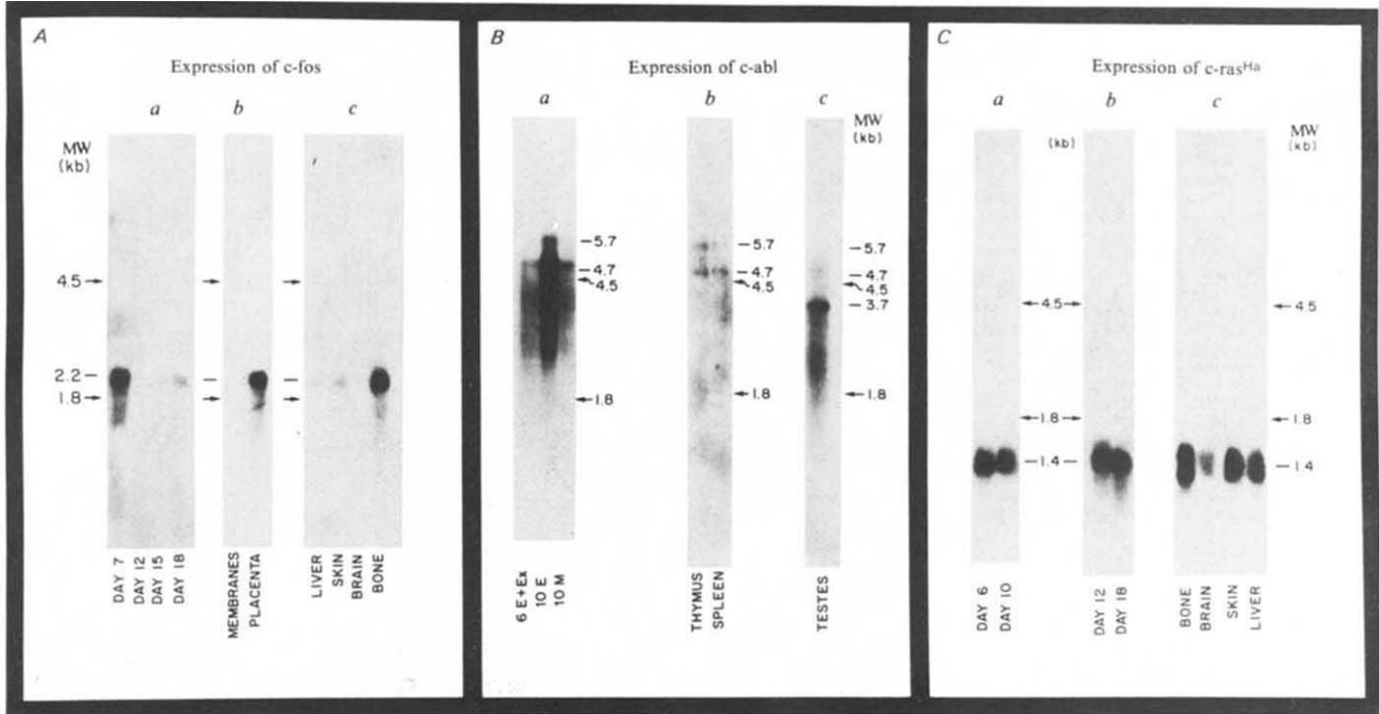


Fig. 3 A, *c-fos* Transcripts in total mouse embryos/fetuses and various tissues. Twenty-five microgrammes of poly(A⁺) RNA (selected once; a, c) or 10 µg of poly(A⁺) RNA (selected twice; b) (isolated as described in Fig. 1a legend) were denatured in the presence of 1 M glyoxal, 20 mM sodium phosphate pH 7.0, for 1 h at 50 °C. The denatured RNA was separated on horizontal 1.1% agarose gels made up and run in 10 mM sodium phosphate buffer pH 7.0 for about 20 h at 35 V, and thereafter blotted on to nitrocellulose paper with 20 × SSC36. Blots were baked for 3 h at 80 °C and prehybridized, hybridized and washed as described for dot blots in Fig. 1a, except that the hybridization was carried out in the presence of 5 × 10⁶ c.p.m. of nick-translated probe per ml for 60 h at 45 °C. Blots were exposed for 60 h (a) and 18 h (b, c), respectively. a, Day 7 embryos containing extraembryonal tissues; day 12, 15, 18 total embryos/fetuses. b, Day 10 membranes and day 10 placenta. c, Bone and skin from 1–3-day-old mice; brain and liver from 10-day-old mice. Arrows indicate the positions of 18S/28S ribosomal RNA markers. B, *c-abl* Transcripts in total mouse embryos and tissues. Ten microgrammes of poly(A⁺) RNA (selected once; a) or 20 µg of poly(A⁺) RNA (selected twice; b, c) were electrophoresed and analysed as described above. Blots were exposed for 24 h (a, c) and 96 h (b), respectively. a, Total embryos and extraembryonal tissue (see Fig. 2a legend). b, c, Postnatal mouse tissues (see Fig. 2b legend). c, *c-ras^{Ha}* Transcripts in mouse embryos/fetuses and various postnatal mouse tissues. Ten microgrammes of poly(A⁺) RNA (selected once; a) or 20 µg of poly(A⁺) RNA (selected twice; b, c) were separated by agarose gel electrophoresis and analysed by Northern blotting essentially as described above. Blots were exposed for 24 h. a, b, RNA from day 6, 10, 12 and 18 embryos/fetuses (day 6 embryos included also extraembryonal tissues). c, Various postnatal tissues (see Fig. 2b legend).

findings lend support to the widely held notion that *c-onc* genes are involved in developmental processes like cell proliferation and differentiation^{25,26}.

We thank Drs E. Adamson, A. Berns, T. Curran, T. Hunter, B. Sefton, C. Van Beveren and M. Vogt for critical reading of the manuscript and valuable suggestions, S. Pitts for technical assistance, Drs S. P. Goff, D. R. Lowy and S. M. Tilghman for

recombinant plasmids containing sequences specific for *v-abl*, *v-ras^{Ha}* and AFP respectively and Dr T. Curran for *v-fos*-specific fragment. R.M. is the recipient of a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. This work was supported by grants from NIH and ACS (to I.M.V.), and from Brown and Williamson, Philip-Morris, Inc., R. J. Reynolds, and U.S. Tobacco Co. (to M.J.C.).

Fig. 4 Level of *c-onc* expression in embryos compared with virus-transformed cell lines. 1.5 µg of poly(A⁺) RNA (selected once) from embryos at day 11 of prenatal development and various amounts of poly(A⁺) RNA from virus-transformed cell lines were spotted on nitrocellulose paper and hybridized to *onc*-specific probes as described above. The blot was exposed for 24 h. Relative absorbances indicate the values measured by density scanning of the autoradiogram normalized to 100 for 0.5 µg of RNA. * Absorbance readings obtained from autoradiograms exposed for only 6 h; RAW253 (cell line derived from an A-MuLV-induced mouse B-cell lymphoma)³⁷, ANN-1 (NIH/3T3 cell line transformed by A-MuLV *in vitro*)³⁸, 3T3-Ha8-21 (NIH/3T3 cell line transformed by HaSV *in vitro*)³⁹.

RNA (µg)	RAW 253	Relative absorbance	ANN-1	Relative absorbance	11 Day embryo	Relative absorbance	3T3-Ha8-21	Relative absorbance	11 Day embryo	Relative absorbance
1.5		286*		248*		34		252*		47
0.5		100		100				100		
0.15		42		36				45		
0.05		10		10				10		
0.015		<5		<5				<5		
0.005		<5		<5				<5		

abl

ras^{Ha}

Received 7 June; accepted 12 August 1982.

1. Weiss, A., Teich, N. M., Varmus, H., & Coffin, J. M. (eds) *RNA Tumor Viruses. The Molecular Biology of Tumor Viruses Part III* (Cold Spring Harbor Monograph 10C, Cold Spring Harbor, New York 1982).
2. Bishop, J. M. *Scient. Am.* **246**, 68–78 (1982).
3. Coffin, J. M. *et al. J. Virol.* **40**, 953–957 (1981).
4. Bishop, J. M. *et al. Rev. Biochem.* **47**, 35–88 (1978).
5. Chen, J. H. *J. Virol.* **36**, 162–170 (1980).
6. Shibuya, M., Hanafusa, H. & Balducci, P. L. *J. Virol.* **42**, 143–152 (1982).
7. Gonda, T. J., Sheiness, D. K. & Bishop, J. M. *Molec. cell. Biol.* **2**, 617–624 (1982).
8. Finkle, M. P., Biskis, B. O. & Jenkins, P. B. *Science* **151**, 698–701 (1966).
9. Finkel, M. P. & Biskis, B. O. *Prog. exp. Tumor Res.* **10**, 72–111 (1968).
10. Curran, T. & Teich, N. M. *J. Virol.* **42**, 114–122 (1982).
11. Curran, T., Peters, G., Van Beveren, C., Teich, N. M. & Verma, I. M. *J. Virol.* (in the press).
12. Abelson, H. T. & Rabstein, L. S. *Cancer Res.* **30**, 2208–2212 (1970).
13. Cook, W. D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2917–2921 (1982).
14. Witte, O. N. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2488–2492 (1978).
15. Witte, O. N., Rosenberg, N. E. & Baltimore, D. *Nature* **281**, 396–398 (1979).
16. Harvey, J. J. *Nature* **204**, 1104–1105 (1964).
17. DeFeo, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3328–3332 (1981).
18. Langbeheim, H., Shih, T. Y. & Scolnick, E. M. *Virology* **106**, 292–300 (1980).
19. Scolnick, E. M. *et al. Molec. cell. Biol.* **1**, 66–74 (1981).
20. Moloney, J. B. *Natn. Cancer Inst. Monogr.* **22**, 139–142 (1966).
21. Frankel, A. E. & Fischinger, P. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3705–3709 (1976).
22. Gattoni, S. *et al. Molec. cell. Biol.* **2**, 42–51 (1981).
23. Kirby, D. R. S. in *Early Conceptus, Normal and Abnormal* (ed. Park, W. W.) 68–73 (University of St Andrews Press, Edinburgh, 1965).
24. Slamon, D. J., Müller, R., Cline, M. J. & Verma, I. M. *Science* (submitted).
25. Comings, D. E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3324–3328 (1973).
26. Bishop, J. M. *ICN-UCLA Symp. molec. cell. Biol.* **23**, 515–524 (1981).
27. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
28. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
29. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
30. Jones, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2651–2655 (1980).
31. Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. *Cell* **22**, 777–785 (1980).
32. Ellis, R. W. *et al. J. Virol.* **36**, 408–420 (1980).
33. Tilghman, S. M. *et al. J. biol. Chem.* **254**, 7393–7399 (1979).
34. Wilson, J. R. & Zimmerman, E. F. *Dev. Biol.* **54**, 187–200 (1976).
35. Dziadek, M. & Adamson, E. J. *Embryol. exp. Morphol.* **143**, 289–313 (1978).
36. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
37. Raschke, W. *Cold Spring Harb Symp. quant. Biol.* **44**, 1187–1194 (1979).
38. Scher, C. D. & Siegler, R. *Nature* **253**, 729–731 (1975).
39. Maisel, J., Scolnick, E. M. & Duesberg, P. H. *J. Virol.* **16**, 749–753 (1975).

Ia invariant chain detected on lymphocyte surfaces by monoclonal antibody

Norbert Koch, Susanne Koch
& Günter J. Hämmerling

Institut für Immunologie und Genetik,
Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280,
D-6900 Heidelberg, FRG

The immune response (I) region of the major histocompatibility complex encodes a set of cell-surface glycoproteins, the Ia antigens, which are probably identical with the products of immune response genes. Chemically, the polymorphic Ia antigens controlled by the I-A and I-E subregions consist of two associated polypeptide chains, α and β (ref. 1). In addition, this α - β complex is noncovalently associated with another chain of 31,000 molecular weight which, because of its non-polymorphic nature, has been designated invariant (I_i) chain². Whereas Ia antigens are known to span the plasma membrane³, the I_i chain has been found in the cytoplasm but not on the surface of lymphocytes². Here we describe a monoclonal antibody detecting free I_i chains on the cell surface which are not associated with Ia antigens.

The belief that I_i chains were not expressed on the cell surface was based on the finding that I_i chains are only observed when metabolically labelled antigens are immunoprecipitated with anti-Ia antibodies and not when surface-iodinated material was investigated⁴. In an attempt to raise monoclonal antibodies against constant parts of murine histocompatibility antigens we have obtained a rat anti-mouse hybridoma, In-1. In a cellular radioimmunoassay, In-1 reacted only with Ia-positive B-cell tumours and spleen cells, but not with Ia-negative T- or B-cell tumours (Table 1). In-1 was positive on spleen cells of all H-2 haplotypes tested so far (H-2^{b,d,f,k,i,r,s}), indicating the non-poly-

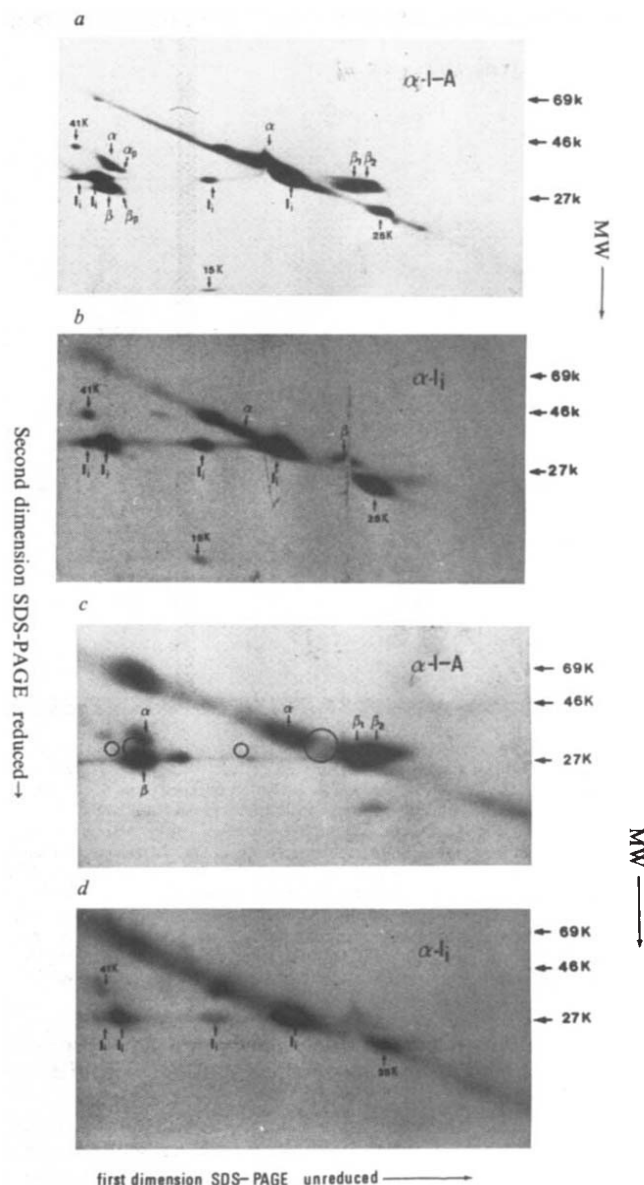


Fig. 1 Two-dimensional SDS-polyacrylamide gel electrophoresis (PAGE) (unreduced-reduced) of Ia and I_i immunoprecipitates. *a, b*, Fluorographs of metabolically (³⁵S-methionine) labelled Ia antigens. 10⁷ blast cells were collected on day 3 after lipopolysaccharide stimulation and labelled for 1 h in the presence of 200 μ Ci ³⁵S-methionine. The *Lens culinaris* binding material was immunoprecipitated with: *a*, monoclonal anti-I-A^k antibody 17/227⁵ or *b*, with In-1 followed by rabbit anti-rat immunoglobulin. Immune complexes were isolated via protein A Sepharose. *c, d*, Autoradiographs of ¹²⁵I surface-labelled (lactoperoxidase) C57BL/6 cells. *c*, *Lens culinaris* binding material was immunoprecipitated using monoclonal anti-I-A^b antibody K25-8.7 (ref. 18). Circles indicate positions where I_i chains usually appear in this type of gel system. This becomes evident when *c* and *d* are superimposed. *d*, C57BL/6 spleen cells were surface labelled with ¹²⁵I. To avoid competition by cold cytoplasmic I_i which is present in excess¹⁰, iodinated cells were incubated with In-1 for 1.5 h. Excess antibody was removed by washing and cells solubilized with 1% NP40. Immune complexes were isolated with rabbit anti-rat immunoglobulin followed by protein A Sepharose.

morphic nature of the determinant recognized. Our failure to block the binding of In-1 to spleen cells by addition of 12 different monoclonal anti-Ia antibodies (refs 5, 18), or by a mixture of these anti-Ia antibodies suggested that the respective determinant did not reside on an Ia antigen.

By immune precipitation of ³⁵S-methionine-labelled spleen cells and separation in one-dimensional 10% SDS gel, a band

with an apparent molecular weight of 31,000 (31K) was observed (not shown). From all these findings, the molecular weight, the non-polymorphic nature and the expression on Ia-positive cells, it seemed that this glycoprotein may be identical with the I_i chain although it is expressed on cell surfaces.

To prove this hypothesis, we used a two-dimensional gel system which allows identification of complexes and their respective subunits⁶. The precipitates are first separated in unreduced form and then, after reduction, in the second dimension. In this type of gel subunits of previously S-S-linked complexes are located in the dimeric region exactly one above each other, whereas components pre-existing as monomers are found on the diagonal. Using this technique we have recently shown that I_i chains occur in at least three different dimeric forms⁷. A typical example of two-dimensional gel analysis of ³⁵S-labelled Ia antigens immunoprecipitated from CBA spleens with monoclonal anti-I-A^k antibody 17/227 (ref. 5) is presented in Fig. 1a. In the dimeric region subunits of the following dimers can be observed (from left to right): (1) I_i plus a component of molecular weight 41K, (2) an I_i homo-dimer, (3) α - β dimers and dimers consisting of their precursors, $\alpha_p\beta_p$, and (4) I_i plus a component of about 15K. Previous alkylation studies have demonstrated that disulphide-linked α - β dimers are formed during solubilization of cells^{7,8}, whereas I_i dimers seem to pre-exist as S-S-linked complexes⁷. The monomeric region shows the characteristic Ia α -polypeptide, single I_i chains, two different forms of β chains which have been described previously⁹, and another component of about 25K molecular weight⁹. The important aspect is that the I_i chain can be identified in three different and characteristic dimeric forms. Two-dimensional peptide analysis demonstrates that the 41K, 25K and 15K components do not share peptides and that they are not structurally related to α or β chains (manuscript in preparation). Figure 1b shows ³⁵S-labelled material immunoprecipitated with In-1. The most prominent spots in Fig. 1b are identical with those for monomeric and dimeric forms of I_i . In addition, traces of α and β chains can be seen in the monomeric but not in the dimeric region. These findings strongly suggest that the antigen recognized by In-1 is identical with the I_i chain. The presence of only small amounts of spots migrating like I-A α and β chains (Fig. 1b) indicates that not all I_i chains, which are in excess in the cytoplasm¹⁰, are associated with Ia. Additional evidence that In-1 is indeed directed against I_i stems from a cell-free translation system in which In-1 precipitated exclusively I_i chains but no other products (J. Lipp *et al.*, manuscript in preparation).

Table 1 Binding of In-1 to Ia positive cells

Cell type	Ref.	H-2	Surface Ia	Surface Ig	Binding of In-1 (c.p.m.)
CH1 tumours	11	a	+	+	3,357 ± 128
WEHI 231 tumours	12	d	+	+	1,319 ± 79
38C13 tumours	13	k	-	+	130 ± 25
EL4 tumours	14	b	-	-	184 ± 47
TL ⁺ tumours	15	k	-	-	167 ± 34
Ag8.653 tumours	16	d	-	±	88 ± 19
CBA spleen		k	+	+	2,805 ± 135

Monoclonal antibody In-1 was obtained by fusion of x63 Ag8.653 myeloma cells and splenocytes from Wistar rats immunized with lymphocytes of strain A.TL followed by injections with B10.A-derived CH1 tumour cells using standard fusion procedures⁵. Hybridoma supernatants were screened on Ia-positive (CH1) and Ia-negative (38C13) B-cell tumours in a cellular binding assay with 20,000 c.p.m. iodinated sheep anti-rat immunoglobulin as a developing reagent which was preabsorbed with mouse immunoglobulin. Data show c.p.m. (± s.e.) bound per 5×10^5 cells. Background with irrelevant monoclonal rat antibody (100–200 c.p.m.) has been subtracted. Ia positivity of CH1 and WEHI 231 was confirmed with monoclonal anti-Ia antibodies in cellular binding assays. In-1 is an IgG2b antibody (G. Moldenhauer, unpublished), and it is not cytotoxic.

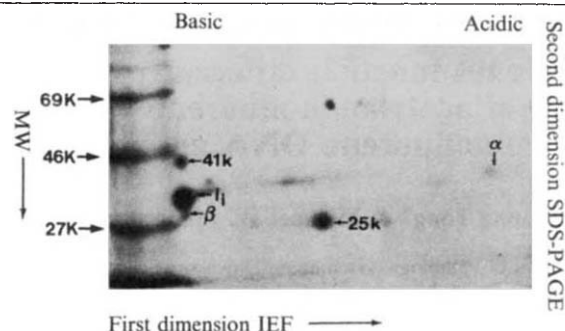


Fig. 2 Two-dimensional IEF gel of I_i immunoprecipitate. Fluorograph of metabolically ³⁵S-methionine labelled Ia antigens. 5×10^6 CH-1 tumour cells were labelled for 1 h in the presence of $100 \mu\text{Ci}$ ³⁵S-methionine. *Lens culinaris* binding material was immunoprecipitated with In-1 and rabbit anti-rat immunoglobulin followed by protein A Sepharose. Two-dimensional IEF gels were performed according to O'Farrell¹⁷.

Because cellular binding assays (Table 1) show that In-1 also reacts with antigens expressed on the cell surface, we immunoprecipitated surface-iodinated molecules. When anti-I-A^k antibodies are used for precipitation, mostly iodinated β chains, some α chains and no invariant chains are seen (Fig. 1c). These data, which are in accordance with published results^{2,4}, have been used to argue that I_i chains are not expressed on the cell surface. In contrast, precipitation of surface-iodinated material with In-1 results in the characteristic I_i spots in the dimeric and monomeric regions (see Fig. 1d). No spots corresponding to α and β chains are visible. Further evidence that the antigen recognized by In-1 on the cell surface is the I_i chain is given by two-dimensional SDS-isoelectrofocusing (IEF) gels (Fig. 2) in which the predominant spot is located in approximately the same position ($p_i \sim 7.5$ –8.0) as has been reported by other investigators for I_i (ref. 4).

In conclusion, we have shown that the monoclonal antibody In-1 binds to a glycoprotein which appears identical to the I_i chain. The two surprising findings were that the I_i chain is expressed on the cell surface and that the cell-surface material is not associated with the Ia α and β chains as is the case in the cytoplasm. The lack of association on the cell surface explains why anti-Ia antibodies have not been suitable reagents for the detection of I_i on cell surfaces.

Whereas it has been speculated that intracellular I_i is involved in transport of Ia α and β chains to the cell membrane¹⁶, the biological function of cell-surface bound I_i is not yet understood. Monoclonal antibody In-1 may be useful in addressing this question and also for isolation and studies on structure and genetics of the I_i polypeptide.

We thank Mr K. Meese for technical help and Mrs A. Riedl for preparation of the manuscript. S.K. is the recipient of a research stipend from the German Cancer Research Centre.

Received 24 March; accepted 3 August 1982.

- Cullen, S. E., Freed, J. H. & Nathanson, S. G. *Transplant Rev.* **30**, 236–270 (1976).
- Jones, P. P., Murphy, D. B., Hewgill, D. & McDevitt, H. D. *Immunochemistry* **16**, 51–60 (1978).
- Walsh, F. S. & Crumpton, M. J. *Nature* **269**, 307–311 (1977).
- Suny, E. & Jones, P. P. *Molec. Immun.* **18**, 899–913 (1981).
- Lemke, H., Hämmerling, G. J. & Hämmerling, U. *Immun. Rev.* **47**, 175–206 (1979).
- Springer, T. A., Kaufmann, J. F., Liddoway, L. A., Mann, D. L. & Strominger, J. L. *J. Biol. Chem.* **252**, 6201–6207 (1977).
- Koch, N. & Hämmerling, G. J. *J. Immun.* **128**, 1155–1158 (1982).
- Cook, R. G., Uhr, J. W., Capra, J. D. & Vitetta, E. S. *J. Immun.* **121**, 2205–2212 (1978).
- Koch, N. & Hämmerling, G. J. *Immunogenetics* **14**, 437–444 (1981).
- Kvist, S., Wiman, K., Claesson, L., Peterson, P. A. & Dobberstein, B. *Cell* **29**, 61–69 (1982).
- Lynes, M. A., Lanier, L. L., Babcock, G. F., Wettstein, P. J. & Haughton, G. J. *Immun.* **121**, 2352–2357 (1978).
- McKean, D. J. *et al. J. exp. Med.* **154**, 1419–1431 (1981).
- Bergmann, Y. & Haimovich, J. *Eur. J. Immun.* **7**, 413–417 (1977).
- Old, L., Boyse, E. & Stockert, E. *Cancer Res.* **25**, 813–819 (1965).
- Ralph, P. J. *Immun.* **110**, 1470–1475 (1973).
- Kearney, J. F., Radbruch, A., Liesegang, B. & Rajewsky, K. *J. Immun.* **123**, 1548–1550 (1979).
- O'Farrell, P. H. *J. Biol. Chem.* **250**, 4007–4021 (1975).
- Koch, N., Hämmerling, G. J., Tada, N., Kimura, S. & Hämmerling, U. *Eur. J. Immun.* (in the press).

uvr Genes function differently in repair of acetylaminofluorene and aminofluorene DNA adducts

Moon-shong Tang* & Michael W. Lieberman

Department of Pathology, Washington University School of Medicine, St Louis, Missouri 63110, USA

Charles M. King

Department of Chemical Carcinogenesis, Michigan Cancer Foundation, Detroit, Michigan 48201, USA

The repair of UV-damaged DNA in *Escherichia coli* is controlled by the *uvr* genes (*A*, *B* and *C*)^{1,2}. These genes are also thought to be involved in the repair of chemically damaged DNA³⁻⁵, however their specific roles in repairing the multiplicity of different adducts that can be produced by chemical carcinogens are not well understood. We have investigated the repair of DNA damaged by different chemical carcinogens and transfected into *E. coli* strains with mutations in different *uvr* genes. The major product of DNA damage by *N*-acetoxy-2-acetylaminofluorene (NA-AAF) is *N*-(guanine-C8-yl)-2-acetylaminofluorene (G-C8-AAF)⁶. Our results suggest that, like UV-induced pyrimidine dimers, the repair of this adduct requires all three *uvr* genes. However, the repair of DNA damaged by *N*-hydroxy-aminofluorene (N-OH-AF, the major adduct of which is *N*-(guanine-C8-yl)-2-aminofluorene, G-C8-AF) seems only to require the *uvrC* gene.

To study the direct biological effect of chemically induced DNA damage and the role of different *uvr* genes in its repair, we have modified Φ X174 RF DNA *in vitro* and used it for transfection of different *E. coli* C strains with a nearly isogenic background (Table 1). This approach is important because chemicals damage a variety of cellular constituents including DNA, RNA, proteins, membranes and small molecules, and the contributions of these various types of damage to toxicity and survival cannot be determined if the whole cell is treated. We have validated our transfection system by studying transfection of wild-type *E. coli* C and the *uvrA*, *uvrB* and *uvrC* strains with UV-irradiated Φ X174 RF DNA. As expected, the transfection frequency of UV-irradiated phage DNA shows the same extent of decrease in all three mutants, as compared with the wild-type cells (Fig. 1a). These three *uvr*⁻ mutants also show significantly lower transfection frequencies than wild-type cells

when NA-AAF-modified Φ X174 RF DNA is used (Fig. 1b).

The lethality of NA-AAF-modified DNA is due mainly to the G-C8-AAF adduct. This conclusion is based on several observations. First, in our conditions, NA-AAF treatment produces at least 85% G-C8-AAF adducts (Fig. 2a). Second, less than two bound carcinogen molecules per molecule of Φ X174 RF DNA are needed to reduce the transfection frequency to 37% (from the Poisson distribution, one lethal hit would reduce survival to 37%, that is, the D_{37}). Thus most of the lethality must result from the G-C8-AAF adduct. In support of this conclusion is the observation that the G-C8-AF adduct is relatively non-lethal (Fig. 1c, see below). These results also demonstrate that all three *uvr* genes are necessary for repair of the G-C8-AAF adduct. The low D_{37} value for G-C8-AAF is similar to that for pyrimidine dimers (Fig. 1a). This result is consistent with the observation that both dimers and this adduct stall DNA replication near the site of DNA damage⁷.

In contrast, treatment of Φ X174 RF DNA with N-OH-AF results in the formation of only the G-C8-AF adduct (Fig. 2b). Transfection studies with this G-C8-AF-containing DNA reveal two distinct features (Fig. 1c): (1) compared with wild-type *E. coli*, two strains carrying independently derived *uvrC* mutations have significantly lower transfection frequencies, while *uvrA* and *uvrB* cells have the same transfection frequency; (2) compared with G-C8-AAF, about 10 times as many G-C8-AF adducts are needed to reduce the transfection frequency to the same extent. Thus, not only is the G-C8-AF adduct less lethal than the G-C8-AAF adduct, but also its repair occurs independently of *uvrA* and *uvrB* gene function. In addition, these results cannot be accounted for by loss of *alk*⁸ function, since both of our *uvrC* strains are *alk*⁺ (methyl methanesulphonate resistant).

These data suggest that all bulky chemical adducts are not repaired in precisely the same way by the *uvr* system. An attractive hypothesis to explain these data relates to steric alterations in DNA. G-C8-AAF lesions assume a *syn* conformation in the DNA helix and result in a distortion which may be similar to that produced by pyrimidine dimers^{9,10}; hence these lesions may have similar requirements for repair. On the other hand, the G-C8-AF products are preferentially in the *anti* conformation and therefore result in little distortion¹¹; this and similar lesions may require only the *uvrC* gene product for increased survival (repair). Thus, the *uvrC* gene product may be required for the nucleotide excision repair of a wide range of DNA adducts while *uvrA* and *uvrB* gene products may be required only for those which produce a great deal of distortion of the DNA helix. Another implication of our findings is that the *uvrC* gene product can function either as a complex with the *uvrA* and *uvrB* gene products (reviewed in ref. 2) or as an independent entity. In this light it is interesting to consider the relationship of the seven known complementation groups of

* Present address: The University of Texas System Cancer Center, Science Park—Research Division, Smithville, Texas 78957, USA.

Table 1 The bacterial strains used

Strain designation	Repair associated marker	Other markers	Source
MST1	Wt	<i>thr</i> -1, <i>leu</i> -6, <i>proA</i> 2, <i>his</i> -4, <i>argE</i> 3 or <i>arg</i> -49, <i>lacY</i> 1, <i>galK</i> 2, <i>rpsL</i> 31, or <i>rpsL</i> 154, <i>supE</i> 44 An <i>E. coli</i> K-12 × <i>E. coli</i> C hybrid ϕ X sensitive	HF4714, T. Kunkel
MST3	<i>uvrB</i> 5	Same as MST1 except <i>Gal</i> ⁺	P1::Tn9c·SR289×HF4714 (select <i>Gal</i> ⁺)
MST8	<i>uvrC</i> 34	Same as MST1 except <i>His</i> ⁺	T4GT7·MST5×HF4714 (select <i>His</i> ⁺)
MST13	<i>uvrA</i> 6	Same as MST1 except <i>Arg</i> ⁺	P1::Tn9c·SR114×MST11 (select <i>Mal</i> ⁺)
MST14	<i>uvrC</i> 56	Same as MST1 except <i>His</i> ⁺	T4GT7·N17-7×MST1 (select <i>His</i> ⁺)
SR289	<i>uvrB</i> 5	<i>leuB</i> 19, <i>metE</i> 70, <i>thyA</i> 36, <i>deo</i> (2?), <i>lacZ</i> 53, <i>rha</i> -5, <i>rpsL</i> 151	W3110, N. Sargentini
SR114	<i>uvrA</i> 6	<i>thr</i> -1, <i>leuB</i> 6, <i>proA</i> 2, <i>his</i> -4, <i>argE</i> 3, <i>lacY</i> 1, <i>mtl</i> -1, <i>galK</i> 2, <i>xyl</i> -5, <i>thi</i> -1, <i>tsx</i> -33, <i>rpsL</i> 31, <i>ara</i> -14, <i>supE</i> 44	AB1886, N. Sargentini
N17-7	<i>uvrC</i> 56	<i>try</i> , <i>gal</i> , <i>sm</i> ⁺	T. Kato
MST5	<i>uvrC</i> 34	Same as SR231 except <i>His</i> ⁺	P1 _{kc} ·SR550×SR231 (select <i>His</i> ⁺)
SR231	<i>uvrC</i> 34	Same as SR114	N. Sargentini
SR550	Wt	Same as SR289	N. Sargentini
MST11	Wt	Same as MST1 except <i>Arg</i> ⁺ , <i>malB</i> 45	This study

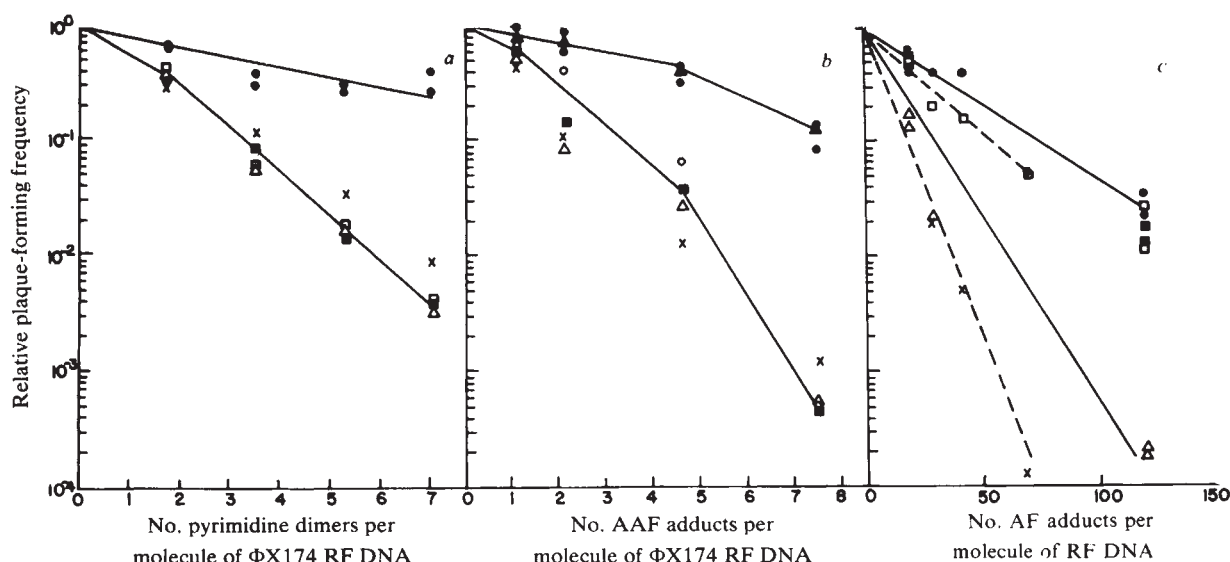


Fig. 1 Relative transfection frequency of Φ X174 am3cs70 RF DNA damaged with; a, UV; b, NA-AAF; c N-OH-AF in different *E. coli* C strains. Φ X174 am3cs70 RF DNA was prepared using the method described by Barnes¹⁵. Different *E. coli* *uvr*⁻ markers ($\square$, *uvrA6*; $\blacksquare$, *uvrB5*; $\triangle$, *uvrC34*; $\times$, *uvrC56*; are introduced into the same *E. coli* C $\times$ K12 hybrid (HF4714) parental strain ($\bullet$) by P1 or T4 transduction (Table 1). *E. coli* C *uvrA* 103 strain ($\circ$) was obtained from Drs T. Kunkel and L. Loeb, and *E. coli* C wild-type strain ($\blacktriangle$) from Dr D. Berg. Cells were grown to $2-5 \times 10^8$ cells ml⁻¹ in Luria broth (LB; 10 g tryptone, 5 g yeast extract and 15 g NaCl in 1 l of H₂O) with 0.1% glucose and 50 μ g ml⁻¹ thymine. The cells were spun down and resuspended in cold 50 mM CaCl₂ and stored at 4 °C overnight. They were concentrated to about 9×10^9 cells ml⁻¹ by centrifugation before transfection. Transfection was carried out by incubating 0.1 ml of chilled DNA in 10 mM Tris, pH 7.5, and 1 mM EDTA with 0.2 ml of CaCl₂-treated cell suspension in an ice water bath for 20 min. At the end of incubation, 3 ml of LB soft agar (6 g agar in 1 l LB broth with 1 mM CaCl₂) were added to the mixture and plated on LB plates (12 g agar in 1 l LB broth). The plates were incubated at 37 °C for at least 5 h before counting. The transfection frequency for control DNA ranged from 10^{-5} to 10^{-6} . The relative transfection frequency was calculated by the formula N/N_0 , where N is the number of plaques formed per unit damaged DNA and N_0 the number of plaques formed per unit of control DNA. ³H-(ring-labelled)NA-AAF (in dimethyl sulphoxide) or ³H-N-OH-AF (in ethanol) was reacted with DNA in 10 mM sodium citrate (for N-OH-AF) or sodium acetate (for NA-AAF) at pH 5.5. The unbound labelled carcinogens were removed by multiple extractions with ether, dialysis and precipitation with ethanol (NA-AAF), or by successive extractions with phenol and ether and precipitation with ethanol (N-OH-AF). The number of adducts formed per molecule of Φ X174 RF DNA was calculated from the ³H specific activity of NA-AAF (710 mCi mmol⁻¹) or N-OH-AF (193 or 206 mCi mmol⁻¹) used for treatment and the specific activity of the DNA. The dotted and solid lines in c represent DNA modified by different batches of ³H-N-OH-AF. The number of pyrimidine dimers formed per molecule of Φ X174 RF DNA was calculated using a dimer yield of 2.6×10^{-5} dimers per thymine residue per J per m² (ref. 16).

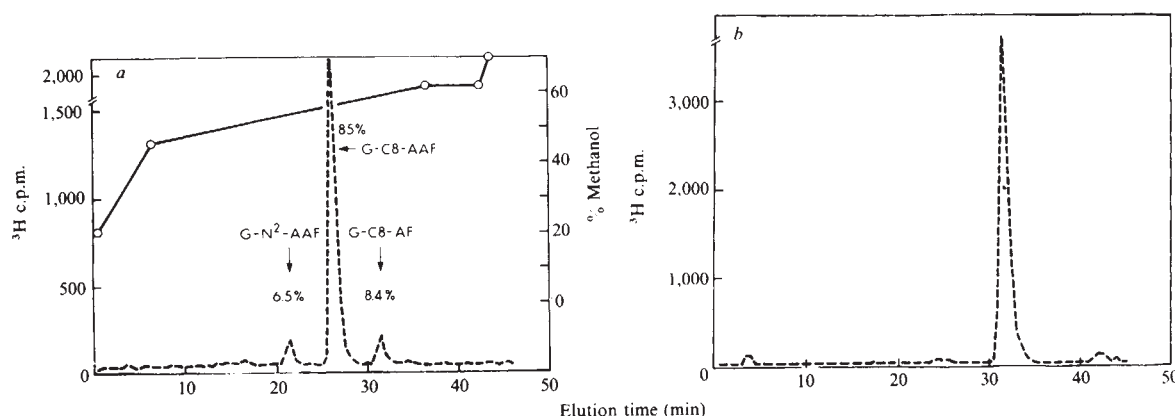


Fig. 2 Chemical analysis of DNA adducts formed in Φ X174 RF DNA after modification by; a, ³H-NA-AAF; or b, ³H-N-OH-AF. NA-AAF-modified Φ X174 RF DNA was hydrolysed by digestion with DNase I, snake venom phosphodiesterase and alkaline phosphatase¹³ and subsequent treatment with trifluoroacetic acid (TFA) at 70 °C for 30 min (M.S.T. and M.W.L., in preparation). The DNA modified with N-OH-AF was treated with TFA at 70 °C for 30 min directly. The hydrolysed mixture was dried under nitrogen and reconstituted in 30% methanol for HPLC. A 0.1 ml sample was injected into a Varian 8500 HPLC with 10 μ m Alltech-C18 column (300 $\times$ 4 mm). The adducts were eluted with a gradient of 21–70% methanol in 8 mM (NH₄)₂PO₄ (pH 3.8; 20 °C) over a 48 min period at a flow rate of 1 ml min⁻¹. Standards (dG-C8-AAF and dG-C8-AF) were added to the mixture as internal markers before adding TFA. A dG-N²-AAF standard was treated with TFA and chromatographed separately, and its position is indicated in the profile. TFA treatment removes deoxyribose from the adducts¹⁷. 0.5 ml samples were counted. The solid line represents the methanol concentration for the elution.

xeroderma pigmentosum in human cells to the repair of chemical damage in these cells.

One plausible explanation for the low lethality of G-C8-AF is that cells tolerate this kind of DNA damage well. In animals or in mammalian cell culture, G-C8-AF has been found to remain in DNA for a long period of time^{12,13}. However, N-OH-AF has been demonstrated to be a very potent mutagen in the Ames' *Salmonella* mutation test¹⁴. These two findings suggest that DNA replication or transcription machinery may be able to read through G-C8-AF without stalling but with poor fidelity.

We thank Dr Frederick A. Beland for the gift of dG-C8-AAF and dG-N²-AAF adducts. This work has been supported by

NIH grants CA 09118 and ES CA 02740 (M.S.T.), CA 20513 (M.W.L.) and CA 23386 (C.M.K.), and by the following companies: Brown & Williamson Tobacco Corporation, Philip Morris Incorporated, R. J. Reynolds Tobacco Company and United States Tobacco Company.

Received March 2; accepted July 26 1982.

- Howard-Flanders, P., Boyce, R. P. & Theriot, L. *Genetics* **53**, 1119–1136 (1966).
- Seeberg, E. *Prog. Nucleic Acid Res. molec. Biol.* **26**, 217–226 (1981).
- Venitt, S. & Tarmy, E. M. *Biochim. biophys. Acta* **287**, 38–51 (1972).
- Ikenaga, M., Ichikawa-Ryo, H. & Kondo, S. *J. molec. Biol.* **92**, 341–356 (1975).
- Hanawalt, P. C., Cooper, P. K., Ganesan, A. K. & Smith, C. A. *Rev. Biochem.* **48**, 783–836 (1979).
- Miller, E. C. & Miller, J. A. *Cancer* **47**, 2327–2345 (1981).

7. Moore, P., Bose, K. K., Rabkin, S. D. & Strauss, B. S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 110-114 (1981).
8. Yamamoto, Y., Katsuki, M., Sekiguchi, M. & Otsuji, N. *J. Bact.* **135**, 144-152 (1978).
9. Trifonov, E. N., Shafraunskaya, V. N., Frank-Kamenetskii, M. D. & Lazurkin, Yu. S. *Molec. Biol.* **2**, 698-705 (1968).
10. Fuchs, R. P. & Daunc, M. P. *Biochemistry* **13**, 4435-4440 (1974).
11. Leng, M., Ptak, M. & Rio, P. *Biochem. biophys. Res. Commun.* **96**, 1095-1102 (1980).
12. Visser, A. & Westra, J. G. *Carcinogenesis* **2**, 737-740 (1981).
13. Howard, P. C., Casciano, D. A., Beland, F. A. & Shaddock, J. G. Jr *Carcinogenesis* **2**, 97-102 (1981).
14. Ames, B. N., Gurney, E. G., Miller, J. A. & Bartsch, H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3128-3132 (1972).
15. Barnes, W. M. *J. molec. Biol.* **119**, 83-99 (1978).
16. Howard-Flanders, P. A. *Rev. Biochem.* **37**, 175-200 (1968).
17. Westra, J. G. & Visser, A. *Cancer Lett.* **8**, 155-162 (1979).

A protein immunologically related to erythrocyte band 4.1 is found on stress fibres of non-erythroid cells

Carl M. Cohen*†, Susan F. Foley* & Catherine Korsgren*

* Department of Research, Division of Hematology and Oncology, St Elizabeth's Hospital, Boston, Massachusetts 02135, USA

† Departments of Medicine and Molecular Biology, Tufts University School of Medicine, Boston, Massachusetts 02111, USA

Erythrocyte band 4.1 is one of the three major structural proteins of the erythrocyte cytoskeleton, along with spectrin and actin¹⁻³. When bound to the tail-end of the actin-binding protein spectrin, band 4.1 is capable of promoting the formation of stable ternary complexes with spectrin and F-actin⁴⁻⁷, both in solution and on the erythrocyte membrane⁸. Recent work has demonstrated that proteins immunologically and structurally related to spectrin, as well as ankyrin—spectrin's membrane attachment site—are present in a wide variety of non-erythroid cells⁹⁻¹³, although the function of these proteins in such cells is still unclear. It has been reported¹⁴ that band 4.1 is present in platelets and granulocytes. Here we show that a protein immunologically related to band 4.1 is present in a non-erythroid cell, the fibroblast, and appears by immunofluorescent labelling to be distributed uniformly along stress fibres. We suggest that by analogy with erythrocyte band 4.1 this protein may be an accessory actin-binding protein, and may serve to integrate actin filaments and actin-binding proteins.

Antibodies to electrophoretically pure band 4.1 were produced in New Zealand white rabbits by multiple intradermal injection of 60 µg freshly prepared^{4,15} band 4.1. Antisera were tested for reactivity against band 4.1 by immunoperoxidase staining of erythrocyte ghost proteins that had been electrophoresed in a 5-15% acrylamide gel, and transferred electrophoretically to nitrocellulose paper. Anti-band 4.1 IgG was purified by affinity chromatography of serum IgG on a band 4.1 affinity column. Figure 1 shows that reaction of affinity purified anti-4.1 with erythrocyte membrane proteins gave one major band in the position of band 4.1 as well as several minor bands of lower molecular weight (see below) (Fig. 1, lane b).

Although it cannot easily be seen in Fig. 1, in the Laemmli buffer system¹⁶ band 4.1 migrates as a closely spaced doublet (4.1a and b), which has recently been shown by peptide mapping to consist of related phosphoproteins¹⁷. Other electroblots (not shown) clearly demonstrate that our antibody is equally reactive with these two components of band 4.1.

To demonstrate further the identity of the peroxidase stained band with band 4.1, we briefly prestained a separate nitrocellulose strip with Coomassie blue, so that the protein bands could be visualized, and then counterstained by the immunoperoxidase procedure. A distinct brown band appeared exactly on top of the Coomassie blue stained band 4.1, verifying the identity of the reactive protein (not shown). Pretreatment of ghosts or intact erythrocytes with diisopropyl fluorophosphate (DFP), pepstatin, phenylmethyl sulphonyl fluoride (PMSF) or EDTA had no effect on the appearance of the lower molecular weight bands, which are likely to be membrane-

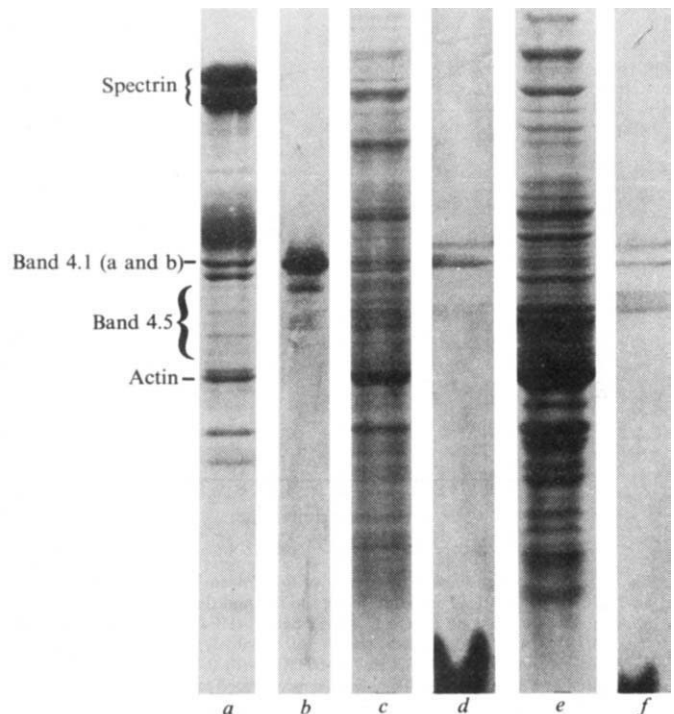


Fig. 1 a, Erythrocyte membrane proteins electrophoresed on a 5-15% acrylamide gradient gel¹⁶ and stained with Coomassie blue. b, Immunoperoxidase stained electroblot of ghost proteins. Erythrocyte membrane proteins electrophoresed in 1.5 mm thick Laemmli slab gels were transferred to nitrocellulose paper (Schleicher & Schuell BA83)²⁴ in a Bio-Rad transblot apparatus at 30 V for 18 h, with cooling. Electroblots were stained with anti-band 4.1 antibody and counterstained with peroxidase conjugated second antibody as described in ref. 25. Coomassie blue staining of the nitrocellulose paper, as well as of the residual proteins on the electroblotted gel, showed that all proteins were transferred with at least 50% efficiency. Affinity purified anti-band 4.1 IgG was prepared from rabbits as follows. Three New Zealand white rabbits received 20-40 subcutaneous injections of purified band 4.1 (total of 60 µg per rabbit) emulsified with an equal volume of Freund's complete adjuvant. After 7 weeks, each of the rabbits produced sera which were positive for anti-band 4.1 activity by the above assay. None of the rabbits' preimmune sera showed any reactivity against red cell membrane proteins. IgG was prepared from sera by two cycles of precipitation at 4 °C with 37% saturated ammonium sulphate. The final precipitate was resuspended in phosphate-buffered saline (PBS) and dialysed for 48 h against 5 mM Na phosphate pH 7.0. After clarification at 32,000g for 15 min the material was passed down a column of Whatman DE 52 equilibrated with 5 mM Na phosphate pH 7.0, and the void volume collected and concentrated by precipitation with 50% saturated ammonium sulphate pH 7.8. The concentrated IgG was dialysed for 48 h against 8 l of ice cold PBS. Electrophoretically pure band 4.1 was coupled to Bio Rad Affigel 10 in 50 mM HEPES pH 7.6 and rabbit IgG was passed over the column and flushed with PBS. Anti-band 4.1 IgG was eluted from the column with 0.2 M glycine pH 2.8 and the fractions were collected into tubes containing sufficient 1 M Na phosphate pH 7.4 for neutralizations²⁶. After recording the A₂₈₀ of the fractions they were stabilized by the addition of 0.02% w/v NaN₃ plus 0.25 mg ml⁻¹ bovine serum albumin and dialysed against PBS. c, High speed pellet of homogenized fibroblasts. Fibroblasts were grown on glass coverslips to about 80% confluency in Dulbecco's minimal essential medium (MEM) plus 10% fetal calf serum and removed by treatment with 0.25% trypsin for 10 min. The cells were washed in PBS containing 2 mM PMSF and suspended to a density of 1 × 10⁷ ml⁻¹ in 0.34 M sucrose, 5 mM EGTA, 10 mM dithiothreitol DTT, 0.5 mM ATP, 2 mM PMSF and 20 mM imidazole-HCl pH 7.0²⁵. The cells were then homogenized in a Teflon-glass homogenizer until >90% of the nuclei were released, as monitored by phase contrast microscopy. The suspension was centrifuged at 140g for 6 min to remove nuclei and unlysed cells and the supernatant was spun at 17,500 r.p.m. for 20 min in a Sorvall SS34 rotor. The resulting pellet (c, d) was solubilized directly into SDS gel solution while the supernatant (e, f) was precipitated with 10% w/v trichloroacetic acid before solubilization. Gel lane c contains ~40 µg of protein. d, Anti-band 4.1 immunoperoxidase staining of fibroblast proteins shown in c transferred to nitrocellulose paper as described for b except that the electroblotting time was increased to 24 h. The staining procedure is identical to that described for erythrocyte membranes except these fibroblast strips were stained with 45 µg affinity purified anti-band 4.1 IgG in 3 ml for 16 h at 37 °C. The lane shown originally contained a total of 170 µg protein. e, Coomassie blue stained gel of high speed supernatant obtained from fibroblast homogenate described in c 60 µg protein. f, Anti-band 4.1 immunoperoxidase stain of supernatant proteins shown in lane e. Electroblotting and staining were done exactly as described in d. The lane shown originally contained 84 µg protein.

bound proteolytic fragments of band 4.1. Purified band 4.1 showed the same pattern of secondary bands as did ghosts. All of these secondary bands lined up with Coomassie blue stained bands visible in the band 4.5 region¹⁸ of ghosts.

Band 4.1 extracted by our standard method^{4,15}, or by an entirely unrelated method using Tween 20¹⁹, appeared homogeneous on overloaded Fairbanks gels¹⁸, but on Laemmli gels revealed a small band (~5% of total protein) migrating just above band 4.1 in addition to the low molecular weight fragments. This band also appeared in ghosts stained by the peroxidase method (Fig. 1, lane *b*), but did not line up with any visible Coomassie blue stained band. While the relationship of this band to band 4.1 is still unclear, preliminary studies

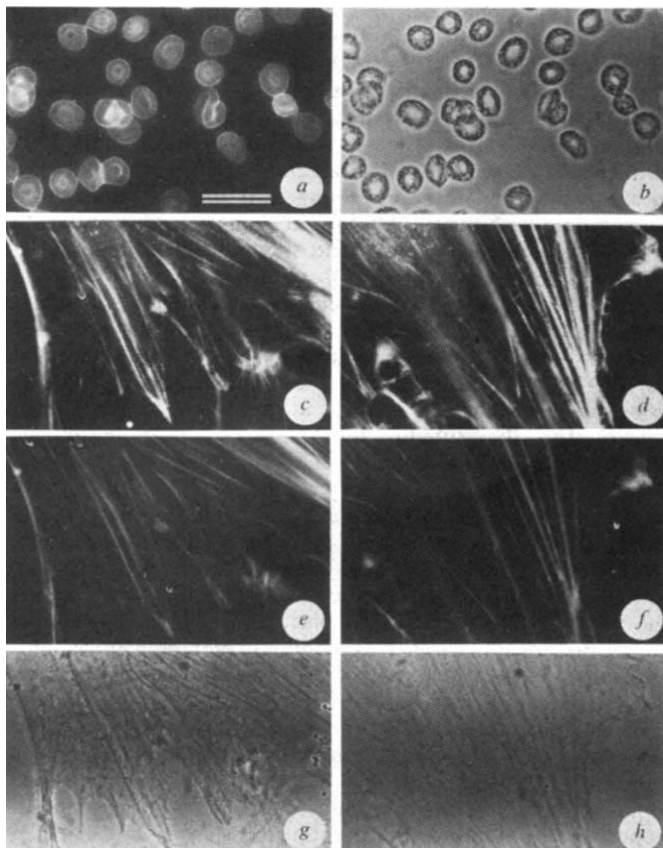


Fig. 2 Phase and immunofluorescent micrographs of fibroblasts and erythrocytes. *a*, Human erythrocytes were washed and diluted in PBS and smeared on a glass slide and allowed to dry. The cells were fixed for 20 min at room temperature in 3.7% formaldehyde in PBS, washed in PBS and dehydrated in acetone at -10°C for 10 min. After rinsing in PBS the fixed cells were incubated with $40\ \mu\text{l}$ of $30\ \mu\text{g ml}^{-1}$ affinity purified anti-band 4.1 antiserum in PBS plus $0.25\ \text{mg ml}^{-1}$ bovine serum albumin for 30 min at 37°C . The slides were rinsed for 30 min in PBS and then $40\ \mu\text{l}$ of 1:30 diluted rhodamine conjugated goat anti-rabbit IgG (Cappel, $5\ \text{mg ml}^{-1}$) was added and incubated for 30 min at 37°C . Slides were washed in 10 changes of PBS over 1 h and coverslips were placed on a drop of 90% glycerol, 10% PBS and sealed with wax. Slides were examined in a Leitz Dialux 20 microscope equipped with a Ploem epifluorescent attachment and a $\times 100$, 1.32 numerical aperture NPL Fluotar lens. Rhodamine fluorescence was viewed with a Leitz N2 fluorescence cube. Scale bar, $2\ \mu\text{m}$. *b*, Erythrocytes from *a* viewed in phase contrast optics. *c, d*, Portion of fibroblasts after staining with affinity purified anti-band 4.1. Fibroblasts were grown as described in Fig. 1 legend on glass coverslips, and were fixed and dehydrated as described above except that they were not allowed to air dry. Coverslips were stained as described for erythrocytes by inverting them on to a $40\text{-}\mu\text{l}$ drop of primary or secondary antibody except that the primary antibody was used at a concentration of $100\ \mu\text{g ml}^{-1}$. Fibroblast in *e, f* were also stained simultaneously for F-actin using NBD-phalloidin. After incubation of cells with primary antibody as described above, the cells were washed and incubated with rhodamine conjugated goat anti-rabbit IgG plus NBD-phalloidin. The phalloidin was added by drying $15\ \mu\text{l}$ of a $2\ \text{ml}$ ($150\ \text{U ml}^{-1}$) stock solution (Molecular Probes, Inc.) in methanol in a glass tube and adding $60\ \mu\text{l}$ of a 1:30 dilution of rhodamine conjugated goat anti-rabbit IgG in PBS. $40\ \mu\text{l}$ of this mixture was then used to stain cells by the procedure described above. *e, f*, F-actin of same cells shown in *c* and *d* viewed by NBD-phalloidin fluorescence using the Leitz I2 fluorescence cube. *g, h*, Phase contrast views of the same cells shown in *2c* and *d* and *2e* and *f* respectively.

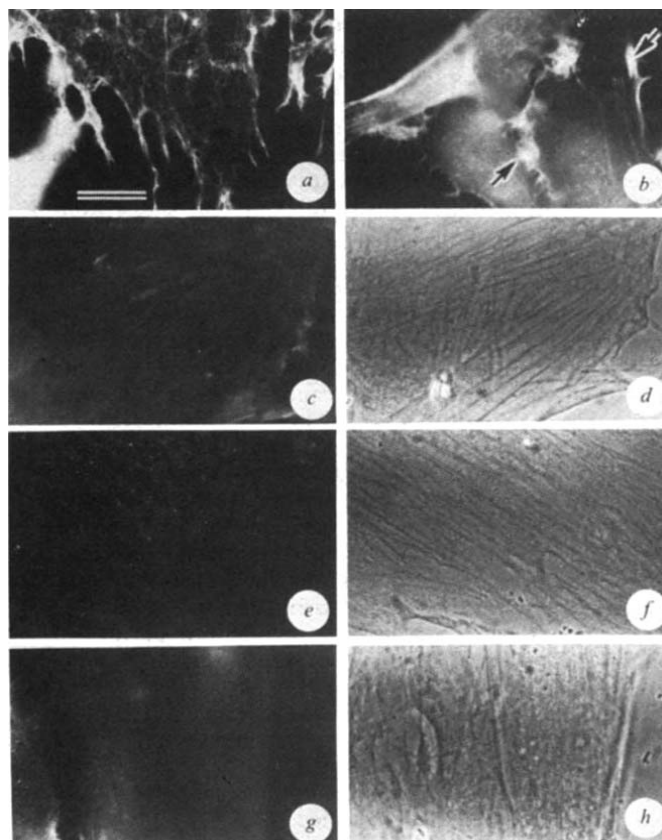


Fig. 3 *a*, Anti-band 4.1 staining of fibroblasts pretreated with $10\ \mu\text{g ml}^{-1}$ cytochalasin B for 60 min at 37°C before fixation and extraction. Cells were treated with affinity purified anti 4.1 as described in Fig. 2 legend. Scale bar, $2\ \mu\text{m}$. *b*, Anti-band 4.1 staining of small poorly spread fibroblasts. Note intense staining in pseudopods and lamellipodia (arrows) but lack of stress fibre staining. *c*, Fibroblasts stained as described in Fig. 2 legend except that $100\ \mu\text{g ml}^{-1}$ normal rabbit IgG was substituted for anti-band 4.1 IgG. *d*, Phase contrast picture of the same cell as in *c* illustrating abundance of stress fibres which were not stained in *c*. *e*, Fibroblasts were stained with $50\ \mu\text{g ml}^{-1}$ anti-band 4.1, washed in PBS and incubated with $40\ \mu\text{l}$ of $100\ \mu\text{g ml}^{-1}$ purified erythrocyte band 4.1 at 37°C for 15 min. The coverslips were washed in PBS and stained with second antibody as described above. *f*, Phase contrast picture of cell in *e*. *g*, Fibroblasts were fixed in formaldehyde but not extracted in acetone, and labelled with primary and secondary antibodies as described for Fig. 2. *h*, Phase contrast picture of cell shown in *g*.

suggest that it is closely related to band 4.1 as it binds to spectrin (unpublished data).

To determine whether band 4.1 was present in a non-erythroid cell we assayed human fibroblasts (ATCC CRL 1139 En Son) grown in culture on glass coverslips, by both immunoperoxidase and immunofluorescence. For the immunoperoxidase method we disrupted fibroblasts in a Dounce homogenizer and after removing nuclei by low speed centrifugation examined the pellet and supernatant produced by subsequent high speed centrifugation for band 4.1. Figure 1, lanes *c, e*, show the Coomassie blue stained gels of the high speed pellet and supernatant, while lanes *d* and *f* show the respective electroblots of these proteins stained by the anti-band 4.1-peroxidase procedure. Both *d* and *f* show a clear peroxidase-stained band which co-migrates exactly with erythrocyte band 4.1, and also a band that migrates just above band 4.1, coincidentally with a similar faint band found in peroxidase-stained erythrocyte membranes mentioned above.

To localize band 4.1 within the cells we examined both fibroblasts and erythrocytes by indirect immunofluorescence using affinity purified anti-band 4.1 IgG. Figure 2*a* shows that staining with anti-band 4.1 IgG resulted in uniform labelling of human erythrocytes. Figure 2*c* and *d* show anti-band 4.1 staining of fibroblasts. The staining appears primarily as bright linear bundles extending to the cell periphery. To determine whether these bundles were identical to stress fibres, we simultaneously stained F-actin in the same cells (Fig. 2*e, f*) using

nitrobenzo oxadiazole (NBD)-phalloidin^{20,21}. Comparison of the band 4.1 and F-actin staining revealed a surprisingly complete correspondence between the two patterns. Further, both sets of patterns coincided well with the appearance of phase dense stress fibres in extracted cells (Fig. 2g, h). Band 4.1 staining seemed particularly intense in ruffles at the leading and upper edge of 'triangular' fibroblasts (Fig. 3b).

To correlate further the distribution of anti-band 4.1 fluorescence with actin, we pretreated the fibroblasts with 10 $\mu\text{g ml}^{-1}$ cytochalasin B before staining. Cytochalasin is known to have profound morphological effects on cells, and leads to the breakdown and disorganization of stress fibres and actin filaments (reviewed in ref. 22). After cytochalasin treatment the cells retracted and began to round up on the coverslips. Figure 3a shows that the pattern of anti-band 4.1 fluorescence was no longer that of linear bundles, but was diffuse and rather disorganized, consistent with the appearance of actin (not shown). Small poorly spread fibroblasts had few or no prominent stress fibres²³, and in these same cells anti-band 4.1 fluorescence was also diffuse (Fig. 3b). Many of these smaller cells had areas of bright anti-band 4.1 staining at their periphery, consistent with band 4.1 being concentrated in lamellapodia or ruffles, as has been reported for actin²³ (arrows, Fig. 3b).

To demonstrate specificity of staining we performed the following controls. (1) Preimmune serum, or non-immune IgG from normal rabbits, gave no reaction with either erythrocyte or fibroblast proteins by the immunoperoxidase method. (2) Pooled normal rabbit IgG gave no fluorescent staining of fibroblasts when substituted for anti-band 4.1 IgG (Fig. 3c). (3) Anti-band 4.1 fluorescence could be displaced or blocked by incubating fibroblasts with purified band 4.1 either during or after antibody labelling (Fig. 3e). (4) No appreciable surface staining was obtained when fixed but non-acetone extracted fibroblasts were reacted with anti-band 4.1 (Fig. 3g).

We conclude that a protein immunologically related to erythrocyte band 4.1 is uniformly distributed along actin filaments in a non-erythroid cell. If the fibroblast protein serves a function similar to erythrocyte band 4.1, then it is possible that it may connect actin filaments to actin-binding proteins and thus to other cellular structures such as plasma or intracellular membranes. Further, by analogy with its interaction with erythrocyte proteins, this protein may be importantly involved in maintaining the elasticity of extended cytoskeletal structures⁷, allowing F-actin-actin binding protein contacts to be made and broken repeatedly without irreversible or long-term cytoskeletal breakdown.

This work was supported by USPHS grant HL 24382 to C.M.C. We thank Dr Deborah K. Smith for many helpful suggestions during this work.

Received 1 July; accepted 23 August 1982.

1. Branton, D., Cohen, C. M. & Tyler, J. *Cell* **24**, 24–30 (1981).
2. Gratzner, W. B. *Biochem. J.* **198**, 1–8 (1981).
3. Sheetz, M. P. *Biochim. biophys. Acta* **557**, 122–134 (1979).
4. Tyler, J., Hargreaves, W. & Branton, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5192–5196 (1979).
5. Ungewickell, E., Bennett, P. M., Calvert, R., Ohanian, V. & Gratzner, W. B. *Nature* **280**, 811–814 (1979).
6. Fowler, V. & Taylor, D. L. *J. Cell Biol.* **85**, 361–376 (1980).
7. Cohen, C. M. & Korsgren, C. *Biochem. biophys. Res. Commun.* **97**, 1429–1435 (1980).
8. Cohen, C. M. & Foley, S. F. *Biochem. biophys. Acta* **688**, 691–701 (1982).
9. Goodman, S. R., Zagan, J. S. & Kulikowski, C. R. (1981) *Proc. natn. Acad. Sci. U.S.A.* **78**, 7570–7574.
10. Davis, J. & Bennett, V. *J. biol. Chem.* **257**, 5816–5820 (1982).
11. Bennett, V. *Nature* **281**, 597–599 (1979).
12. Bennett, V. & Davis, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7550–7554 (1981).
13. Glenney, J. R., Glenney, P., Osborn, M. & Weber, K. *Cell* **28**, 843–854 (1982).
14. Spiegel, J. E., Beardsley, D. S. & Lux, S. E. *Fedn Proc.* **41**, 657 (1982).
15. Cohen, C. M. & Foley, S. F. *J. Cell Biol.* **86**, 694–698 (1980).
16. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
17. Goodman, S. R., Yu, J., Whitfield, C. F., Culp, E. N. & Posnak, E. J. *J. biol. Chem.* **257**, 4564–4569 (1982).
18. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. *Biochemistry* **13**, 2606–2617 (1971).
19. Liljas, L., Lundahl, P., Hjerten, S. *Biochim. biophys. Acta* **352**, 327–337 (1974).
20. Barak, L. S., Yocum, R., Nothnagel, E. A. & Webb, W. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 980–984 (1980).
21. Barak, L. S., Yocum, R. R. & Webb, W. W. *J. Cell Biol.* **89**, 368–372 (1981).
22. Pollard, T. D. & Weihing, R. R. *CRC Crit. Rev. Biochem.* **2**, 1–65 (1974).
23. Herman, I. M., Crisone, N. & Pollard, T. D. *J. Cell Biol.* **90**, 84–91 (1981).
24. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
25. Yin, H., Albrecht, J. H. & Fattoum, A. *J. Cell Biol.* **91**, 901–906 (1981).
26. Fujiwara, K. & Pollard, T. D. *J. Cell Biol.* **71**, 848–875 (1976).

Post-transcriptional control of tubulin biosynthesis during leishmanial differentiation

Michael Wallach, Dunne Fong & Kwang-Poo Chang*

The Rockefeller University, New York, New York 10021, USA

Controls of gene expression can occur at different levels during cell differentiation. Using an *in vitro* culture system¹, we have studied protein biosynthesis during the differentiation of a parasitic protozoan, *Leishmania mexicana*. There are two developmental stages in the life cycle of this and similar parasites: motile extracellular promastigotes in the sandfly gut and non-motile intracellular amastigotes in the mammalian macrophages. Intracellular parasitism by the latter results in human leishmaniasis, diseases still widespread in many parts of the world². During leishmanial differentiation from one stage to the other, we observed changes in tubulin biosynthesis, concomitant with the morphological changes of these parasites in the length of their flagella and subpellicular microtubules³. With few exceptions^{4,5} tubulin biosynthesis is known to be under transcriptional control in other eukaryotic systems^{6–12}. Here we have examined the protein synthesis during leishmanial differentiation in *in vitro* translation systems using total cellular poly(A)⁺ RNA from amastigotes and promastigotes. Contrary to our expectations, the results indicate a post-transcriptional control for the tubulin biosynthesis during leishmanial differentiation in our host-parasite culture system.

Total cellular poly(A)⁺ RNA was extracted from the promastigotes by phenol-SDS extraction¹³ and oligo(dT)-cellulose elution¹⁴. RNA was translated in wheat germ¹⁵ and rabbit reticulocyte lysate¹⁶ cell-free systems and the products were compared with *in vivo* metabolically labelled proteins. As shown in Fig. 1, the 55,000-molecular weight (MW) doublet bands, previously identified as tubulins in metabolically labelled parasites³, were also found in the translation products from poly(A)⁺ RNA. Little or no tubulin was translated *in vitro* using promastigote poly(A)⁺ RNA (not shown). The labelling intensity of tubulin bands seemed to vary with the Mg²⁺ concentration. The lower band of the doublet (β -tubulin) had a higher labelling intensity in sub-optimal conditions than the upper one (α -tubulin). Preferential translation of β - over α -tubulin *in vitro* has also been reported with *Tetrahymena* poly(A)⁺ RNA (ref. 17). However, note that β -tubulin contains more methionine residues than α -tubulin¹⁸. It is also possible that mRNAs compete for the available factors in an *in vitro* translation system^{19,20}. In our study with *Leishmania* RNA, we selected the reticulocyte system for further experiments because of its lower background translation activities.

Our previous finding of a pronounced decrease in the tubulin biosynthesis during leishmanial promastigote to amastigote differentiation led us to determine whether there might be a concomitant decrease in the level of translatable tubulin mRNA. Cell-free translation products of promastigote and amastigote poly(A)⁺ RNAs were compared with each other and with the *in vivo* metabolically labelled proteins from the same batch of parasites (Fig. 2). Tubulins were found, as expected, to be the most prominent bands in metabolically labelled promastigotes (a) and in *in vitro* products from promastigote mRNA (b), but not in metabolically labelled amastigotes (c); an unexpected finding was that mRNA from amastigotes translated *in vitro* (c, d) just as much tubulin as that from promastigotes (a, b). Identical results were obtained when the concentrations of mRNA from both stages were varied from 0.1 to 2 μg for the *in vitro* translation experiments (not shown), thereby eliminating mRNA competition as a possible explanation for the results observed.

*To whom correspondence should be addressed.

As the purity of isolated amastigotes has been verified¹, macrophage RNA could not possibly contribute to the *in vitro* translation products observed. To ascertain the identity and quantity of tubulin translated *in vitro* we performed immunoprecipitation of samples using monospecific rabbit anti-serum against β -tubulin of *Chlamydomonas flagella*²¹. β -Tubulin bands of equal intensity were specifically precipitated from translation products of total cellular poly (A)⁺ RNA from both promastigotes and amastigotes (Fig. 3). Thus similar amounts of translatable tubulin mRNA are present in the two developmental stages of *Leishmania mexicana*.

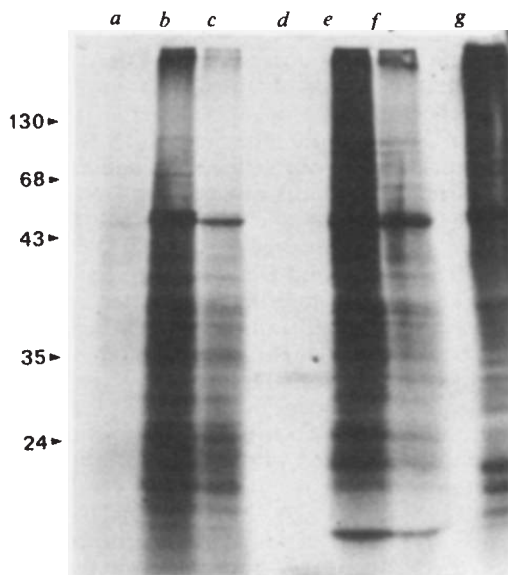
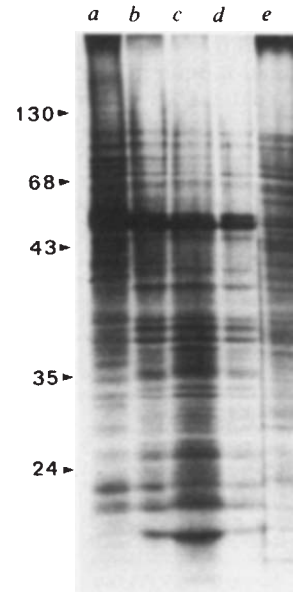


Fig. 1 Translation of promastigote mRNA in the wheat germ and rabbit reticulocyte cell-free systems. *Leishmania mexicana amazonensis* promastigotes were cultured in HEPES buffered Medium 199 plus 10% HIFBS and used after washing three times in Hanks' balanced salt solution by centrifugation. Total RNA was extracted using SDS and phenol¹³ and was chromatographed on an oligo(dT)-cellulose column¹⁴. Total poly(A)⁺ RNA was ethanol precipitated, dried with nitrogen, and dissolved in 10 mM HEPES pH 7.5. Promastigote mRNA was translated in the wheat germ and rabbit reticulocyte cell-free systems (Bethesda Research) as described previously^{15,16}, using ³⁵S-methionine as the labelling substrate. Intact cells were labelled with ³⁵S-methionine for 60 min as previously described³. Total cell-free products were analysed on a 10–15% gradient polyacrylamide gel³. The results using the wheat germ system are as follows: a, no mRNA, 3 mM Mg²⁺; b, promastigote mRNA (1 μ g), 3 mM Mg²⁺; c, promastigote mRNA (1 μ g), 5 mM Mg²⁺. The results using the rabbit reticulocyte system are as follows: d, no mRNA, 3 mM Mg²⁺; e, promastigote mRNA (1 μ g), 3 mM Mg²⁺; f, promastigote mRNA (1 μ g), 5 mM Mg²⁺. Lane g shows the proteins metabolically labelled *in vivo* using the same batch of promastigotes. α - And β -tubulin appear as dense bands in between the 43,000 and 68,000 MW markers. Numbers at the left indicate MW marker proteins: β -galactosidase (130,000), bovine serum albumin (68,000), ovalbumin (43,000), pepsin (35,000) and trypsinogen (24,000).

Our results indicate that the large decrease in tubulin biosynthesis which occurs during leishmanial differentiation from the extracellular promastigote stage to the intracellular amastigote stage is not due to a decrease in the amount of translatable tubulin mRNA. Thus, post-transcriptional regulation must be in operation during cell differentiation of this parasite. It is possible that the tubulin half-lives of promastigotes and amastigotes may be different due to a different level of proteolytic activities between the two stages. It has been reported that protease activities are somewhat higher in the amastigotes than in the promastigotes²². However, this difference between the two stages cannot explain their difference in tubulin proteins, as we have pulse-labelled both for 15–60 min and found no difference in the labelling intensity of the tubulin bands (not shown). It is perhaps more likely that tubulin mRNA of amastigotes may be sequestered, thereby unavailable for the protein biosynthetic machinery. It is known that the masked maternal mRNA has a long half-life during the early development of sea urchin and marine snail embryos^{4,5}. In another trypanosomatid *Crithidia*, about 45% of total translationally competent

Fig. 2 Comparison of proteins synthesized *in vitro* by promastigote and amastigote mRNAs with those metabolically labelled *in vivo*. Promastigotes of *Leishmania mexicana* were obtained and total poly(A)⁺ RNA extracted as described in the legend to Fig. 1. Amastigotes were cultured in cells of the macrophage line J774 and purified by Percoll gradient centrifugation; isolated amastigotes remained viable and free from the contamination of macrophage materials¹. Total amastigote poly(A)⁺ RNA was extracted as described in Fig. 1 legend. Proteins synthesized in the rabbit reticulocyte cell-free system using poly(A)⁺ RNA (1 μ g) and metabolically labelled proteins of intact organisms in the presence of ³⁵S-methionine were analysed by SDS-polyacrylamide gel electrophoresis and autoradiography. a, Proteins synthesized by promastigotes *in vivo*. b, Cell-free products directed by promastigote mRNA. c, mRNA from 9-day-old cultures of amastigotes. d, mRNA from 3-week-old cultures of amastigotes. e, The proteins synthesized by amastigotes *in vivo*. Molecular weight markers are indicated as in Fig. 1.

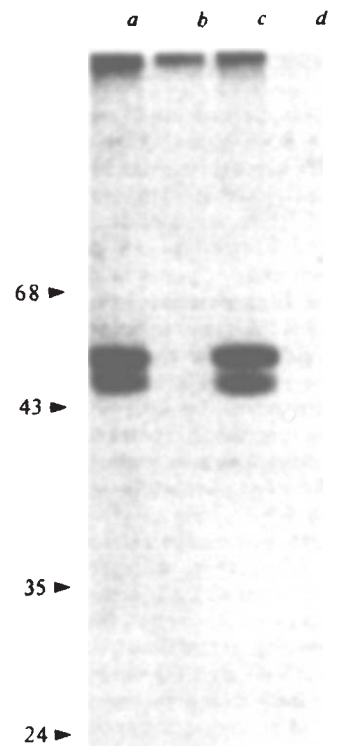


poly(A)⁺ RNA has been found in a non-polysomal compartment²³ and unavailable for translation *in vivo*.

The presence of untranslated tubulin mRNA in amastigotes may have bearings on the evolution of pathogenic *Leishmania* species. Presumably, the promastigote stage in the invertebrates precedes the amastigote stage during the evolution of leishmaniasis²⁴. Retention of tubulin mRNA by amastigotes might be a manifestation of incomplete adaptation of leishmanias in the mammalian host. The presence of abundant tubulin mRNA in the amastigotes could also explain the previous observation that leishmanial differentiation is more rapid from amastigotes to promastigotes than in the reverse direction³.

Thus, we have demonstrated post-transcriptional regulation of tubulin biosynthesis during leishmanial differentiation based on the evidence of *in vitro* translation and β -tubulin

Fig. 3 Immunoprecipitation for β -tubulin of *in vitro* translation products of amastigotes and promastigote mRNAs. Protein samples translated in the reticulocyte system were diluted to about 1.5×10^6 c.p.m. in 50 μ l phosphate-buffered saline (PBS) containing 1% NP40 and incubated for 30 min at 4°C. Each sample was divided into two equal portions; to each was added 10 μ l normal rabbit serum as control or rabbit antiserum against β -tubulin of *Chlamydomonas flagella*²¹. After 18 h at 4°C, each sample received a 50 μ l 33% Protein A-Sepharose 4B in PBS and incubated at 4°C for 2 h under constant agitation. Immunoprecipitates collected in the Protein A beads were extensively washed as previously described³. All samples were subjected to SDS-polyacrylamide gel electrophoresis followed by autoradiography. a and c are immunoprecipitated β -tubulin from amastigote and promastigote samples, respectively. b and d are their respective controls using normal rabbit serum. Note that the 55,000-MW β -tubulin bands immunoprecipitated from amastigotes and promastigote samples are of equal intensity and that there are no protein bands in the controls. The light band below tubulin are artefacts due to the presence of abundant heavy chains of IgG, which is banded at about the same location and brings down some of the radiolabelled tubulins.



immunoprecipitation experiments. This finding can be further studied by using cloned cDNA probes for the analysis of leishmanial tubulin mRNA^{10,11,25}.

We thank Dr Gianni Piperno for the antiserum for β -tubulin of *Chlamydomonas*, Dr A. Kilejian for her support of this project, and Mr J. Stanorski for preparation of this manuscript. M.W. is supported by NIH grant AI-07185 (A.K.); D.F. and K.P.C. are supported by NIH grant AI-15183 and Hirschl Trust (K.P.C.).

Received 21 October 1981; accepted 4 August 1982.

- Chang, K.-P. *Science* **209**, 1240-1242 (1980).
- Zuckerman, A. & Lainson, R. in *Parasitic Protozoa* Vol. 1 (ed. Kreier, J. P.) 57-133 (Academic, New York, 1977).
- Fong, D. & Chang, K.-P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7624-7628 (1981).
- Raff, R. A., Colot, H. V., Selvig, S. E. & Gross, P. R. *Nature* **235**, 211-214 (1972).
- Raff, R. A., Newrock, K. M. & Secrist, R. D. *Devl Biol.* **44**, 369-374 (1975).
- Merlino, G. T., Chamberlain, J. P. & Kleinsmith, L. J. *J. Biol. Chem.* **253**, 7078-7085 (1978).
- Marcaud, L. & Hayes, D. *Eur. J. Biochem.* **98**, 267-273 (1979).
- Lai, E. Y., Walsh, C., Wardell, D. & Fulton, C. *Cell* **17**, 867-878 (1979).
- Lefebvre, P. A., Silflow, C. D., Wieben, E. D. & Rosenbaum, J. L. *Cell* **20**, 469-477 (1980).
- Silflow, C. D. & Rosenbaum, J. L. *Cell* **24**, 81-88 (1981).
- Minami, S. A., Collis, P. S., Young, E. E. & Weeks, D. P. *Cell* **24**, 89-95 (1981).
- Raff, E. C. *et al. Cell* **28**, 33-40 (1982).
- Nienhuis, A. W., Falvey, A. K. & Anderson, W. F. *Meth. Enzym.* **30**, 621-630 (1974).
- Krytosiek, A., Cawthon, M. L. & Kabat, D. *J. Biol. Chem.* **250**, 6077-6084 (1975).
- Roberts, B. E. & Paterson, B. M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2330-2334 (1973).
- Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67**, 247-256 (1976).
- Portier, M. M., Milet, M. & Hayes, D. H. *Eur. J. Biochem.* **97**, 161-168 (1979).
- Kraus, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 4156-4160 (1981).
- Gilmore-Herbert, M. A. & Heywood, S. M. *Biochim. biophys. Acta* **454**, 55-66 (1976).
- Asselbergs, F. A. M., Meulenbergh, E., Van Venrooi, W. J. & Bloemendaal, H. *Eur. J. Biochem.* **109**, 159-165 (1980).
- Piperno, G. & Luck, D. J. *J. Biol. Chem.* **252**, 383-391 (1977).
- Coombs, G. H. *Parasitology* **84**, 149-155 (1982).
- Rondinelli, E., Soares, M. C. M., de Castro, J. F. & de Castro, F. T. *Archs Biochem. Biophys.* **209**, 349-355 (1981).
- Wilson, V. C. L. C. & Southgate, B. A. in *Biology of the Kinetoplastida* Vol. 2 (eds Lumsden, W. H. R. & Evans, D. A.) 241-268 (Academic, London, 1979).
- Cleveland, D. W., Lopata, M. A., Sherline, P. & Kirschner, M. W. *Cell* **25**, 537-546 (1981).

Developmental inactivity of 5S RNA genes persists when chromosomes are cut between genes

J. B. Gurdon*, Colin Dingwall*, Ronald A. Laskey* & Laurence Jay Korn†

* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

† Department of Genetics, Stanford University, Stanford, California 94305, USA

Xenopus and other amphibians possess two major classes of 5S RNA genes¹. One class, which codes for somatic-type 5S RNA, is expressed in both oocytes and somatic cells². The other class, which is 50 times more abundant and encodes oocyte-type 5S RNA, is expressed in oocytes but not in somatic cells²⁻⁴. The oocyte-type 5S genes are located in tandem arrays of hundreds of genes at the ends of most chromosomes⁵, and it has been postulated that the lack of expression of oocyte-type 5S RNA genes in somatic cells is due to some general characteristic of the chromosomal region which the genes occupy⁶, for example, a folding into heterochromatin or higher-order structures. An alternative possibility is that the inactivity of these genes is controlled individually, and is independent of adjacent genetic material. We have now distinguished between these two possibilities by testing the transcription of oocyte-type 5S RNA genes either in their normal arrangement on intact chromosomes or after each gene has been separated from its neighbours by micrococcal nuclease digestion. Transcription was assayed by injecting intact nuclei, containing the 5S RNA genes as chromatin, into *Xenopus* oocytes⁷. We find that the developmental inactivity of oocyte-type 5S RNA genes is retained after the chromatin is cleaved into fragments the size of individual gene repeats.

Micrococcal nuclease was used to cut the interphase chromosomes of intact somatic nuclei. Digestion was arrested by cooling the nuclear suspension on ice or by adding EDTA to 1 mM. (When nuclear suspensions treated in this way are injected directly into oocytes, no further digestion of DNA takes place

in the oocytes, as shown by analysis of re-extracted DNA; data not shown.) After nuclei had been digested with nuclease for varying lengths of time, DNA was extracted and analysed. Digestion for 2 h is sufficient to reduce most of the nuclear DNA to less than 600 base pairs (bp) in length (Fig. 1a). Moreover, transfer of the digested DNA to nitrocellulose and hybridization with [α -³²P]-labelled 5S DNA shows that the chromatin containing 5S RNA genes is digested to the same extent as total chromatin (Fig. 1b). Although this probe will hybridize to both oocyte- and somatic-type 5S RNA genes, the 50-fold excess of oocyte over somatic genes in the *Xenopus* genome⁸ makes it likely that only the oocyte-type 5S RNA genes are seen in this analysis. Hybridization with specific 5S somatic spacer DNA or 5S oocyte spacer DNA gave identical results except that the signal obtained with the somatic spacer was weaker, reflecting the smaller number of somatic genes (data not shown). Because the oocyte-type 5S RNA genes are arranged in repeats of about 650 bp^{9,10}, and because a 2-h digestion of chromatin reduces most of the DNA to less than 600 bp long, there is, on the average, only one 5S RNA gene (120 bp long) in each fragment of chromatin. We estimate that <1% of the 120 min-digested DNA is long enough to contain three genes (Fig. 1b). We note that even after 2 h of nuclease digestion, the nuclei are still intact, and that while the DNA has been cut into small segments, there could conceivably be some continuity of chromosomal fragments as a result of protein interactions.

The transcription of 5S RNA genes was tested by injecting suspensions of nuclei, treated with nuclease for various lengths of time, into non-activating oocytes⁷, which were then labelled for 2 days with [α -³²P]GTP⁷. In these circumstances, when the chromatin is not digested at all, it has been shown that the somatic-type 5S RNA genes are transcribed but the oocyte-type 5S RNA genes in the nuclei remain transcriptionally inactive⁷. (When oocytes from activating females are used the 5S oocyte-type genes are always activated.) Gel analysis of the transcripts (Fig. 2a) shows that there is no activation of the oocyte-type 5S RNA genes at any level of nuclease digestion. Even when the chromosomes have been cut to a length of DNA too small

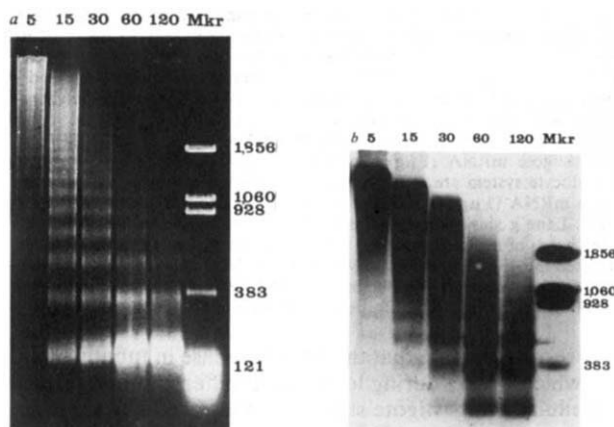


Fig. 1 Agarose gel analysis of DNA extracted from erythrocyte nuclei, after varying times of micrococcal nuclease digestion. *a*, Ethidium bromide stain of total DNA. *b*, Southern transfer of above, hybridized to ³²P-nick-translated pXlo 31, a pBR322 plasmid containing one repeat of *Xenopus laevis* oocyte-type 5S DNA⁹. Lanes marked 5 to 120 are nuclei treated with micrococcal nuclease for 5 to 120 min; Mkr, *Eco*RII digest of pBR322 showing fragment lengths in base pairs. The DNAs used for *b* were from the same samples as for *a*, but were run on a separate agarose gel. *Xenopus* erythrocyte nuclei, prepared by lysolecithin¹⁵, were incubated at 10⁷ nuclei ml⁻¹ in 50 mM Tris-HCl pH 8.0, 1 mM CaCl₂, and 2 U ml⁻¹ of micrococcal nuclease (Worthington Biochemicals) at 37 °C for the times indicated. DNA was prepared from nuclei and analysed on a 1.2% agarose gel. The inefficient transfer of small DNA fragments to nitrocellulose¹⁶ has resulted in a reduced signal from the DNA of mononucleosome length. Note that the 383-bp marker fragment has also transferred inefficiently.

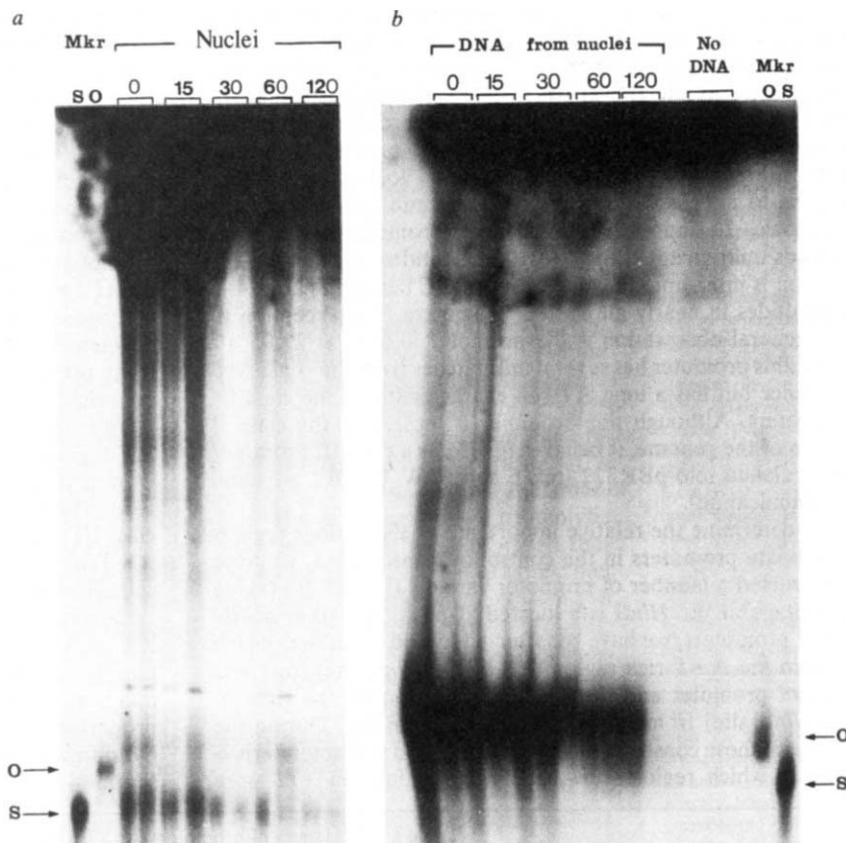


Fig. 2 5S gene transcription from erythrocyte nuclei or DNA injected into *Xenopus* oocytes. *a*, Approximately 500 somatic nuclei (after digestion with micrococcal nuclease for 0–120 min) were injected into individual oocytes, aiming for the germinal vesicle. Preparation of nuclei and nuclease digestions were as for Fig. 1. Oocyte (O) and somatic (S) 5S RNA markers are indicated in the Mkr lanes. *b*, DNA was purified from the same nuclease-treated nuclei as used for *a* and injected into oocytes. All oocytes were subsequently injected with [α - 32 P]GTP and incubated for 2 days. The lane marked 'No DNA' depicts control oocytes not injected with DNA. Labelled RNA was extracted and analysed on non-denaturing polyacrylamide gels⁷. Each track contains the RNA from a single oocyte.

on average to contain more than one gene, the developmental inactivation persists. At each time, the transcription of the somatic-type 5S genes can be seen, though there is less activity at later times. When interpreting this result, it is necessary to recall that there are 50 times more oocyte- than somatic-type genes⁸; a reactivation of oocyte-type genes in only 10% of the injected nuclei would have been readily observed.

One explanation for the non-expression of the oocyte-type 5S genes in nuclease-treated nuclei is that these genes cannot be transcribed as small segments of DNA. Large fragments of 5S DNA injected into oocytes are always transcribed^{11,12}. We have purified DNA from the nuclease-treated nuclei, and injected it into non-activating oocytes. Figure 2*b* shows that oocyte-type 5S RNA is synthesized by DNA from all samples, thus demonstrating that small segments of 5S DNA can be transcribed.

We have shown that the developmental inactivation of 5S RNA genes, once established, can be maintained when chromosomal DNA is cut into pieces containing only one gene. Therefore, although our results do not exclude a need for the tandem arrangement of genes to establish the inactive state initially (see refs 13, 14 for further discussion), they show that the mechanism which maintains the inactive state does not depend on a cooperative effect of many adjacent genes or on distant DNA sequences. Rather, each gene unit must contain regulatory elements sufficient to maintain the inactive state.

We thank S. Whytock for assistance with these experiments, D. Brown for 5S DNA clones, and C. Queen and J. Price for helpful comments on the manuscript. This work was partly supported by NSF grant PCM 8104522 to L.J.K.

Received 23 June; accepted 1 September 1982.

1. Brown, D. D. & Sugimoto, K. *Cold Spring Harb. Symp. quant. Biol.* **38**, 501–505 (1973).
2. Ford, P. J. & Southern, E. M. *Nature new Biol.* **241**, 7–12 (1973).
3. Wegnez, M., Monier, R. & Denis, H. *FEBS Lett.* **25**, 13–20 (1972).
4. Denis, H. & Wegnez, M. *Dev Biol.* **58**, 212–217 (1977).
5. Pardue, M. L., Brown, D. D. & Birnstiel, M. L. *Chromosoma* **42**, 191–203 (1973).
6. Ford, P. J. & Mathieson, T. *Nature* **261**, 433–435 (1976).
7. Korn, L. J. & Gurdon, J. B. *Nature* **289**, 461–465 (1981).
8. Peterson, R. C., Doering, J. L. & Brown, D. D. *Cell* **20**, 131–141 (1980).
9. Federoff, N. V. & Brown, D. D. *Cell* **13**, 701–716 (1978).
10. Carroll, D. & Brown, D. D. *Cell* **7**, 467–475, 477–486 (1976).
11. Brown, D. D. & Gurdon, J. B. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2064–2068 (1977); **75**, 2849–2853 (1978).

12. Gurdon, J. B. & Brown, D. D. *Dev Biol.* **67**, 346–356 (1978).
13. Bogenhagen, D. F., Wormington, W. M. & Brown, D. D. *Cell* **28**, 413–421 (1982).
14. Gottesfeld, J. M. & Bloomer, L. S. *Cell* **28**, 781–791 (1982).
15. Gurdon, J. B. *J. Embryol. exp. Morph.* **36**, 523–540 (1976).
16. Thomas, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

Identification of a potential control region in bacteriophage T7 late promoters

L. K. Jolliffe, A. D. Carter* & W. T. McAllister†

Department of Microbiology, University of Medicine and Dentistry of New Jersey, Rutgers Medical School, PO Box 101, Piscataway, New Jersey 08854, USA

During infection of *Escherichia coli*, bacteriophage T7 encodes a simple RNA polymerase that consists of a single protein species of 883 amino acids^{1,2}. This phage-specified RNA polymerase transcribes the late genes of T7 in two temporal classes (class II and class III)³. The mechanism by which late transcription is regulated is unclear. In general, class II promoters are weaker than class III promoters and *in vitro* the purified phage RNA polymerase discriminates between these classes of promoters as a function of ionic conditions and agents that affect the stability of the DNA helix⁴. Sequence analysis of 17 promoters utilized by the phage RNA polymerase reveals a highly conserved 23 base-pair (bp) consensus sequence that is different from known bacterial promoters⁵. To determine which structural features of the late promoters are important for the regulation of transcription, we have constructed a number of class II/III hybrid promoters and have determined their properties both *in vitro* and *in vivo*. The results reported here indicate that an A + T-rich region that overlaps the upstream end of the promoter has an important role in the control of promoter selection during T7 infection. In addition, changes in the 23-bp conserved promoter sequence may also affect promoter regulation.

* Present address: Laboratory of Biochemistry, National Cancer Institute, Building 37, NIH, Bethesda, Maryland 20205, USA.

† To whom all correspondence should be addressed.

A detailed comparison of promoter structure for the class II and class III promoters (Table 1) reveals a highly conserved region that extends from the +6 position to the -17 position (+1 represents the site at which RNA synthesis is initiated). As a rule, the class III promoters never deviate from the consensus sequence, whereas the class II promoters deviate from the conserved sequence at several locations. Another feature to note is the length of the AT run located between positions -13 and -22. In all class III promoters the AT run extends uninterrupted for 8 to 10 nucleotides. In contrast, the AT run is interrupted by one or more GC base pairs after 4-6 nucleotides in nearly all class II promoters. The exception to this general observation is the promoter located at 27.95 T7 units; this promoter has several mismatches from the conserved sequence but has a long AT run characteristic of the class III promoters. Although this promoter is located in the class II region of the genome, it behaves *in vivo* as a class III promoter when cloned into pBR322 (ref. 6 and F. W. Studier, personal communication).

To determine the relative importance of the various regions of the late promoters in the control of transcription, we have constructed a number of promoter variants (Fig. 1). By taking advantage of the *HinfI* site located at position -10 in nearly all T7 promoters, we have constructed hybrid promoters which contain the A+T-rich region (left of the *HinfI* site) from one class of promoter and the remaining portion (to the right of the *HinfI* site) from the other class of promoter. During the course of these constructions, we also isolated promoter derivatives in which regions downstream from position +12 were

removed from the promoter (pADC10) and in which regions upstream from -23 were removed from the promoter (pLJ1).

To test the behaviour of promoter variants, we first examined their properties *in vitro* by transcribing linearized plasmid DNA that contained the promoters with purified T7 RNA polymerase in the presence of increasing concentrations of KCl. *In vitro*, initiation of transcription by the T7 RNA polymerase at class II promoters in T7 DNA has been observed to be more sensitive to increased ionic strength than is initiation at the class III promoters⁴. As seen in Fig. 2, this is also true of class II and class III promoters that have been removed from T7 DNA and embedded into plasmid DNA sequences. Removal of regions downstream from the +12 position (pADC10) and upstream from position -23 (pLJ1) did not affect the characteristic behaviour of the class III promoters tested.

The class II/III hybrid promoter in pADC5 (which has the AT region from a class II promoter and the initiation region from a class III promoter) behaves as a class II promoter *in vitro*. The class III/II hybrid promoter (in pLJ8) behaves as a class III promoter *in vitro*. These experiments indicate that the A+T-rich region plays a predominant part in determining the salt sensitivity of the late promoters *in vitro*.

During infection of cells carrying recombinant plasmids, T7 late promoters in the plasmids are recognized with similar kinetics as the corresponding late promoters located in the phage genome⁶. This feature makes it possible to examine the properties of the altered promoters *in vivo* by means of a simple hybridization protocol (Fig. 3). In the typical pattern of transcription exhibited by a class II promoter (in pAR95), transcrip-

Table 1 Sequences of promoters for T7 RNA polymerase

Promoter	RNA start site [†]	Nucleotide sequence*
Left end promoter		5' -25 -20 -15 -10 -5 +1 +5 +10 +15 +20 -3'
ϕOL	1.02	TTCTGTTTAAATACAACTCACTATAAGGACAA
Class II promoters		
ϕ1.1A	14.62	TCCGGTTCGGCTAACGCCAAATCAATACGACTCACTATAGGGGACAAACTCAAGTCA
ϕ1.1B	14.81	TTATGATTGACCTTCTTCGGTTAATACGACTCACTATAGGAGACCTTAAGGTTAAC
ϕ1.3	16.02	GACCACTGAAGGACTGGAATCAATACGACTCACTATAGGACAAATGCTTAAGTCCG
ϕ1.5	19.45	CGTGGAGTATAGTTAACTGCTAATACGACTCACTAAAGGAGGTACACCATGATGT
ϕ1.6	19.74	TGATATGCCAGATGCTCAGGCTTAATACGACTCACTAAAGGAGAACTATATGTTCCGA
ϕ2.5	22.77	GCCTCTCTGTAGCAGCGAAGTAATACGACTCACTATAGGGAAC
ϕ3.8	27.95	GTCACTTCTGACCGTGATAATTAATTAAGTCACTAAAGGAGACACACGGG
ϕ4C	31.68	CCGACTGACAAATCCGACTCACTAAAGAGAGACA
ϕ4.3	33.35	GTTTCGTCGACACTCCGATTCGAATCAATACGACTCACTAAAGGAGACACCATGTTCAA
ϕ4.7	34.78	TAAGTGGTGGCTTCATGAATTAATTCGACTCACTATAGGAGATTTACCATGCGTGA
Class III promoters		
ϕ7	46.50	GATAAACTCAAGTCCCTTAATTAATACGACTCACTATAGGAGATAGGGGCTT
ϕ9	54.40	CGAATTAATACGACTCACTATAGGAGACCTCATCTT
ϕ10	57.29	AACTTCTGATAGACTTCGAATTAATACGACTCACTATAGGAGACACACGG
ϕ13	68.00	TCTGGTGGCTCGAAATTAATACGACTCACTATAGGAGACAAATACGACACGG
ϕ17	86.56	AACTTAGGAGCGCTAGGAATTAATACGACTCACTATAGGAGACCGGAATAA
ϕOR	98.24	TTACTGAAAGACACGATAATTAATACGACTCACTATAGGAGACGAGGACGAAAGC
Cloned promoters [‡]		
pAR95	34.78	TAAGTGGTGGCTTCATGAATTAATTCGACTCACTATAGGAGATTTACCATGCGTGA
pADC5	II/III hybrid	TAAGTGGTGGCTTCATGAATTAATTCGACTCACTATAGGAGATAGGGGCTTACCG
pADC10	46.50	GATAAACTCAAGTCCCTTAATTAATACGACTCACTATAGGAGATAGGGGCTTACCG
pLJ8	III/II hybrid	GATAAACTCAAGTCCCTTAATTAATACGACTCACTATAGGAGATTTACCATGCGTGA
pLJ1	68.00	TTTACAGCTTATCATCGAAATTAATACGACTCACTATAGGAGACAAATACGACACGG

Sequences for the promoters at 31.68 and 98.3 T7 units were personal communications from J. Dunn and S. Stahl, respectively. Sequences for the remaining promoters have been published previously¹¹⁻¹⁴. The sequence given for the upstream substitution in pLJ1 is the sequence which would result from cloning a *TaqI* fragment into the *ClaI* site of pBR322 (see Fig. 1).

[†] The location given is in T7 units¹².

[‡] The 23-base pair conserved region is enclosed by the box. Mismatches from the consensus sequence are indicated by an asterisk. The AT run in each promoter is underlined. The dashed lines indicate the upstream and downstream substitutions in the appropriate cloned promoters. The *HinfI* cleavage site is marked by an apostrophe.

[§] Promoter constructions are described in Fig. 1. The hybrid promoters were sequenced by the method of Maxam and Gilbert¹⁵.

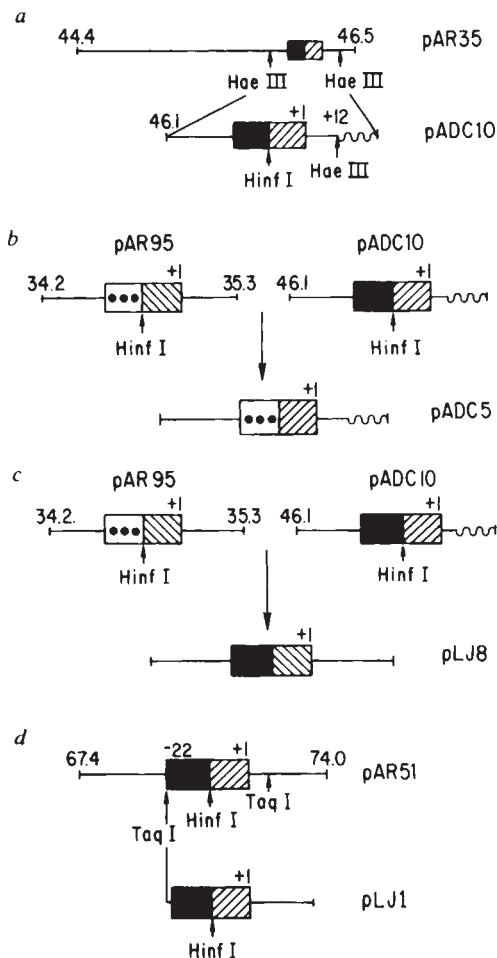


Fig. 1 The constructions described below began with T7 late promoters that had been cloned into pBR322 (ref. 16) using standard procedures¹⁷. Unless otherwise noted, all promoters are inserted into the *Bam*HI site of pBR322. *a*, The class III promoter from 46.5 T7 units (in pAR35; ref. 16) was cloned as a *Hae*III fragment into the *Bam*HI site of pBR322 by ligation of synthetic linkers to give pADC10. During the subcloning, regions downstream from position +12 were substituted by a 51-bp *Hae*III fragment of unknown origin. *b*, The AT run of the class II promoter from 34.7% (in pAR95; ref. 16) was joined to the initiation region of the class III promoter from 46.5% (in pADC10) by ligation at the common *Hinf*I site to obtain a class II/III hybrid promoter (pADC5). *c*, A class III/II hybrid promoter was constructed by joining the left section of the 46.5% promoter (from pADC10) to the right section of the 34.7% promoter (from pAR95) to yield pLJ8. *d*, The upstream flanking regions of the promoter at 68% were removed by digestion of the *Bam*HI insert of pAR51 (ref. 16) with *Taq*I. This fragment was cloned into the *Cla*I site of pBR322. The left portion of class III promoters is indicated by ■; the right portion is denoted by ▨. The left half of a class II promoter is indicated by □; the right half is denoted by ▩. Restriction sites are as indicated. The downstream substitution in pADC10 and pADC5 is indicated by a wavy line. The start site for initiation of RNA synthesis is represented by +1. Sequence analysis of the hybrid promoters confirmed the constructions described above (Table 1).

tion of the plasmid begins early and is markedly reduced by about 10–12 min post infection. In contrast, the class III promoter in pAR35 continues to be used until later in infection. Transcription from the promoter in pADC10 (which has a downstream substitution 12 bp to the right of the initiation site) resembles that of pAR35, indicating that removal of the downstream flanking regions does not affect regulation of this promoter *in vivo*. Similarly, transcription of plasmid sequences controlled by a class III promoter having an upstream substitution 2 bp to the left of the AT run (pLJ1) remains unchanged from that of a class III promoter (that is, pAR35).

The properties of the hybrid promoters were compared with each of the starting promoters by the same techniques (Fig. 3). The behaviour of the class II/III promoter (pADC5) is characteristic of the class II promoter that donated the A+T-rich region (pAR95), whilst the behaviour of the class III/II hybrid promoter (pLJ8) is somewhat intermediate between the class II and class III parent promoters. We conclude from these results that shortening the AT run from 9 to 4 bp is sufficient to cause a class III promoter to be regulated as a class II promoter *in vivo*. However, replacing the short AT run of the class II promoter with the longer AT run of a class III promoter does not seem to overcome completely the effects of changes in the consensus sequence from +2 to +6.

The results of this study enable us to define more precisely the sequences in T7 late promoters that are required for transcription initiation and regulation. Panayotatos and Wells had reported that the region from +33 to –20 bp is sufficient for polymerase binding and initiation *in vitro*⁷. Subsequently, Osterman and Coleman demonstrated that cleavage of the promoter at –10 by *Hinf*I prevented initiation; however, substitution of a nonspecific *Hinf*I fragment of DNA restored initiation of transcription *in vitro*⁸. Although these results suggested that specific T7 sequences to the left of the *Hinf*I site are not required for promoter recognition *in vitro*, note that the non-specific fragments used to reconstitute the promoter in the latter study resulted in the presence of AT regions of 2–4 bp at the same location as in the native promoter sequences. The possibility that this region might participate in promoter function thus remained open.

To answer these questions, we have examined the effects of changes in promoter structure on promoter use *in vivo* and *in vitro*. Our results demonstrate that the region from +13 to –22 bp is sufficient to allow for initiation of transcription both *in vitro* and *in vivo*. Further, we have shown that the sequence

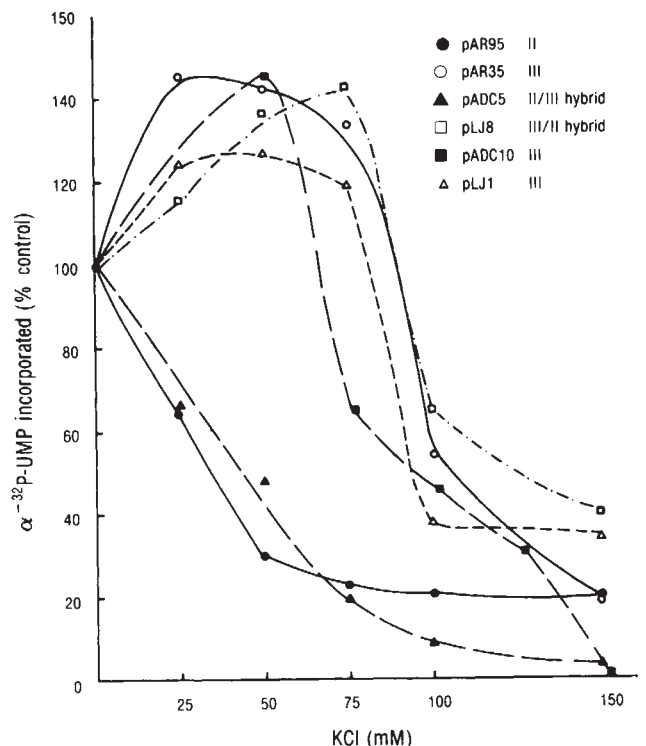


Fig. 2 Plasmids containing T7 promoters linearized by digestion with *Ava*I (which cuts pBR322 only once, and does not cut T7 DNA) were transcribed by purified T7 RNA polymerase in the presence of KCl as indicated. The amount of ³²P-UMP incorporated into acid insoluble products was determined as previously described⁴ and is indicated as per cent of control (in the absence of KCl). The average level of activity in the control samples was 3.1×10^4 c.p.m.

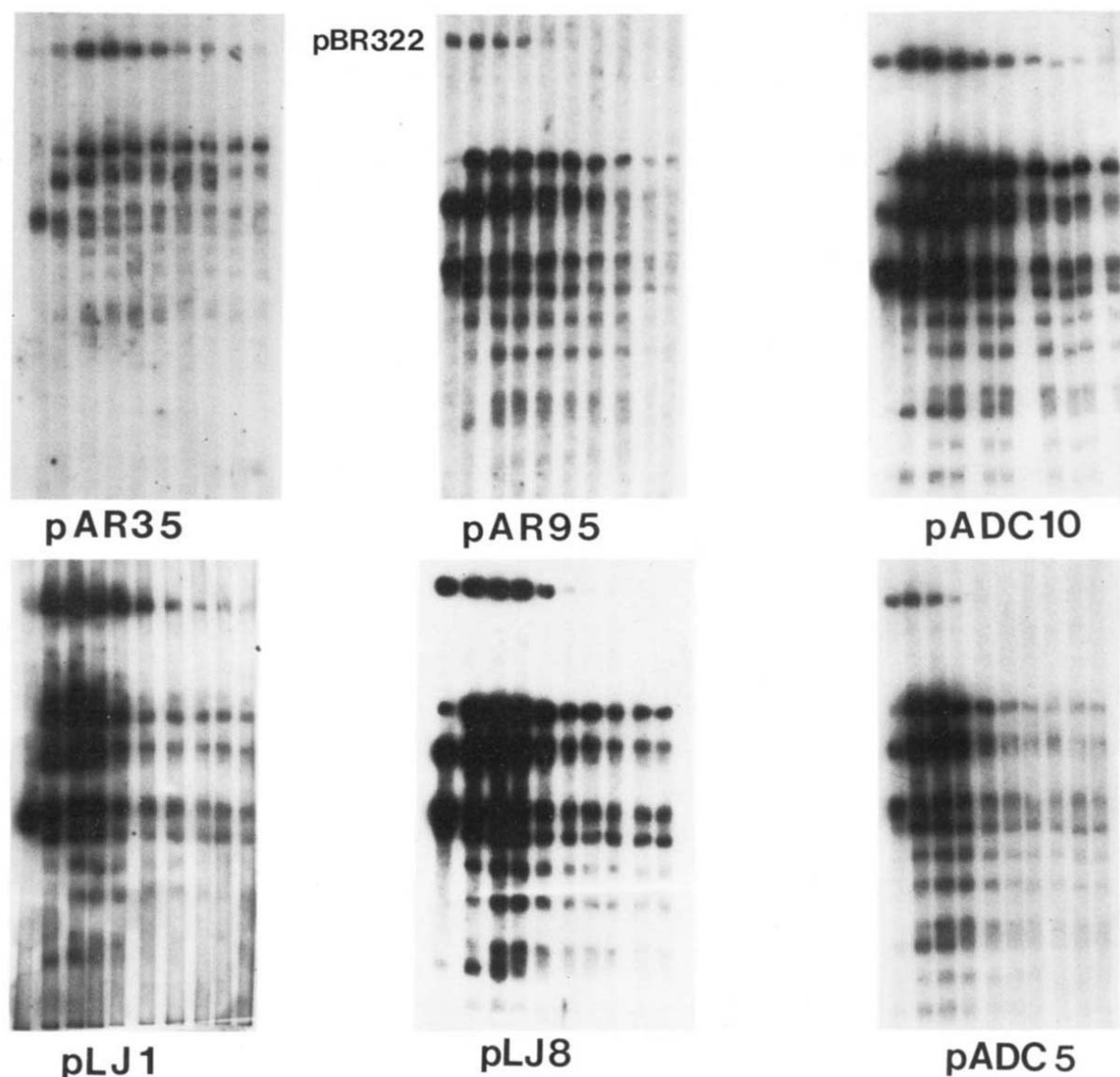


Fig. 3 Bacteria harbouring the plasmids indicated were grown in B2 medium¹⁸ containing thymidine ($20 \mu\text{g ml}^{-1}$) and ampicillin ($20 \mu\text{g ml}^{-1}$), at 37°C to a density of approximately $5 \times 10^8 \text{ cell ml}^{-1}$. The cultures were infected with T7⁺ at a ratio of 10 to 7 phage particles per cell. At intervals following infection, $100 \mu\text{l}$ of culture were removed and added to $\text{H}_3^{32}\text{PO}_4$ ($10 \mu\text{Ci ml}^{-1}$; specific activity, 285 Ci mg^{-1}). Incorporation of label was terminated by the addition of 5 vol. cold $0.01 \text{ mM Na}_2\text{S}_2\text{O}_8$, $75 \text{ mM Na}_2\text{HPO}_4$. The cells were collected by centrifugation and lysed by incubation in lysis buffer (0.02 M sodium acetate, $\text{pH } 5.2$, 1 mM EDTA , $0.5\% \text{ SDS}$). Cell lysates were extracted twice with phenol at 65°C to obtain total RNA⁶. The aqueous phase ($25 \mu\text{l}$) was used directly for hybridization to nitrocellulose filter strips containing *HpaI* fragments of T7 DNA, and pBR322 DNA that had been linearized by digestion with *EcoRI* (ref. 6). Each lane represents successive 2-min pulse intervals beginning at 2 min post infection and ending 22 min post infection. The top band in each panel represents hybridization to plasmid DNA. The bands below the plasmid DNA represent hybridization of labelled RNA to *HpaI* fragments of T7 DNA. Note that the T7 RNA polymerase does not recognize any promoters in pBR322 and, because host transcription is shut off by 4 min after infection, all transcription of plasmid sequences after that time depends on recognition of the T7 late promoter cloned into the plasmid⁶.

in this region provides the necessary information for appropriate regulation of polymerase activity. Whereas the length of the AT run seems to play a primary part in regulating promoter selection *in vitro*, the results are less clear cut *in vivo* and suggest that in some cases changes in the conserved region (+2 to +6) may also be involved in promoter regulation.

Although we have identified a potential control region in late T7 promoters, the mechanism by which promoter selection is regulated *in vivo* remains unclear. We had previously observed that T7 mutants that are defective in gene 3.5 (lysozyme) show an altered transcription programme in which transcription of the class II genes is not shut off with normal kinetics³. It has been reported that expression of gene 3.5 results in changes in the permeability of the cell membrane⁹ and release of phage DNA from a membrane bound complex¹⁰. Whether these changes affect the state of the DNA helix or alter the overall transcriptional capacity of the cell such that only the strongest promoters continue to be used remains to be determined.

This work was supported by NIH grant GM21783 and a grant to W.T.M. from the Foundation of the University of

Medicine and Dentistry of New Jersey (UMDNJ 21-82). L.K.J. and A.D.C. were supported by USPHS Training Grant CA-09069-07. Part of this work was submitted to Rutgers University in fulfillment of the degree requirements of A.D.C.

Received 19 April; accepted 15 July 1982.

1. Chamberlin, M., McGrath, J. & Waskell, L. *Nature* **228**, 227-231 (1970).
2. Stahl, S. J. & Zinn, K. *J. molec. Biol.* **148**, 481-485 (1981).
3. McAllister, W. T. & Wu, H.-L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 804-808 (1978).
4. McAllister, W. T. & Carter, A. D. *Nucleic Acids Res.* **8**, 4821-4837 (1980).
5. Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319-353 (1979).
6. McAllister, W. T., Morris, C., Rosenberg, A. H. & Studier, F. W. *J. molec. Biol.* **153**, 527-544 (1981).
7. Panayotatos, N. & Wells, R. D. *Nature* **280**, 35-39 (1979).
8. Osterman, H. L. & Coleman, J. E. *Biochemistry* **20**, 4884-4892 (1981).
9. Condit, R. C. *J. molec. Biol.* **98**, 45-56 (1975).
10. Silberstein, S., Inouye, M. & Studier, F. W. *J. molec. Biol.* **96**, 1-11 (1975).
11. Carter, A. D. & McAllister, W. T. *J. molec. Biol.* **153**, 825-830 (1981).
12. Dunn, J. J. & Studier, F. W. *J. molec. Biol.* **148**, 303-330 (1981).
13. Rosa, M. D. *Cell* **16**, 815-825 (1979).
14. Oakley, J., Strothkamp, R., Sarris, A. & Coleman, J. *Biochemistry* **18**, 528-537 (1979).
15. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-559 (1980).
16. Studier, F. W. & Rosenberg, A. H. *J. molec. Biol.* **153**, 503-525 (1981).
17. Colowick, S. P. & Kaplan, N. O. *Meth. Enzym.* **68**, 1-555 (1979).
18. Studier, F. W. *J. Bact.* **124**, 307-316 (1975).

MATTERS ARISING

Significance of ankle structures in archosaur phylogeny

IN his recent analysis of archosaur phylogeny¹, Chatterjee places great emphasis on the pattern of articulation between the proximal tarsal bones. He takes the distinction between 'normal' and 'reverse' articulations (Fig. 1*d, e*) to reflect a basic dichotomy of archosaur stocks, in each of which there occurred transitions between mesotarsal (MT) and crurotarsal (CT) joints. These joints are distinguished by the course of the 'hinge' through the ankle (Fig. 1*a, c*). Chatterjee's interpretation (Fig. 2*a*) is not entirely satisfactory: it entails reversals in ankle structure and involves morphological changes that are not readily explained in functional terms.

Chatterjee maintains¹ that the CT joint evolved from the MT joint, and vice versa, even though it is difficult^{2,3} or impossible^{4,5} to envisage such changes. Moreover, Chatterjee suggests that these changes constituted at least two reversals (in parallel) during archosaur history: from MT (in primitive 'sprawling' archosaurs) to CT (in archosaurs with 'semi-improved' stance), and back to MT (in advanced archosaurs with 'erect' or 'fully-improved' stance). There are only three ways in which the MT joint might have evolved from the CT joint (or vice versa): (1) by way of a non-functional ankle in adult animals (with calcaneum locked both to crus and to distal tarsals); (2) by way of a non-functional ankle in embryonic archosaurs; (3) by way of a duplex ankle joint (with 'hinges' of the ankle

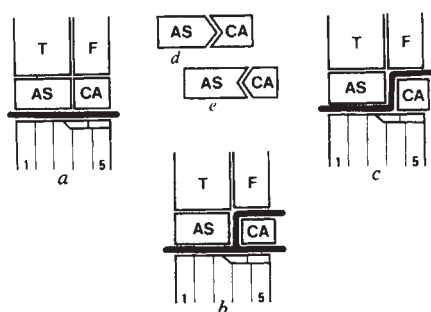


Fig. 1 Block diagrams to illustrate tarsal characters of archosaurs. *a*, Mesotarsal (MT) joint; *b*, duplex (DP) joint; *c*, crurotarsal (CT) joint. In all three cases the left ankle is shown in anterior view, with its principal 'hinge' indicated by a thick line. Bones identified as follows: AS, astragalus; CA, calcaneum; F, fibula; T, tibia; 1, first metatarsal; 5, fifth metatarsal. Both MT and CT joints (*a, c*) might have been derived from the DP joint (*b*) by suppression of appropriate sections of 'hinge' in the latter. *d*, 'Normal' peg-and-socket linkage of proximal tarsals; *e*, 'reverse' peg-and-socket linkage of proximal tarsals.

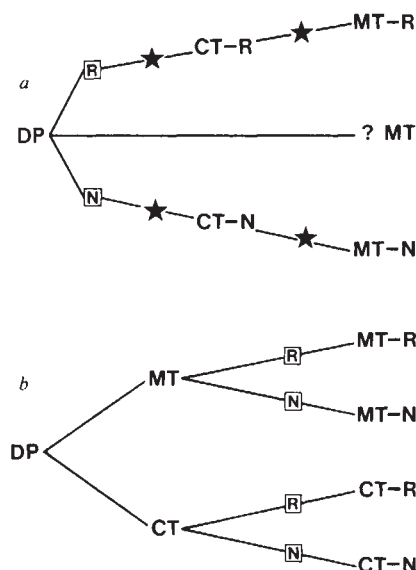


Fig. 2 Distribution of tarsal character-states in *a*, Chatterjee's scheme of archosaur phylogeny¹; and in *b*, alternative scheme proposed here. CT, crurotarsal joint; DP, duplex joint; MT, mesotarsal joint; suffix-N, 'normal' linkage of proximal tarsals; suffix-R, 'reverse' linkage of proximal tarsals; boxed N, development/elaboration of 'normal' linkage; boxed R, development/elaboration of 'reverse' linkage; star, transition from MT joint to CT joint or vice versa.

proximal and distal to the calcaneum; Fig. 1*b*). The first possibility seems unlikely, while the second is untestable. A duplex ankle joint (DP) has been identified in some Triassic archosaurs³, though it has not been regarded as structurally intermediate between MT and CT joints. If the DP joint were intermediate, the scheme proposed by Chatterjee would entail even more complicated reversals: MT → DP → CT → DP → MT.

Peg-and-socket articulation of the proximal tarsals is widespread among Triassic archosaurs^{1,6}, irrespective of the basic ankle joint (whether DP, MT or CT). Vestiges of such articulations ('normal' or 'reverse') persist in the MT ankles of some advanced archosaurs, and rudiments of them ('normal' plus 'reverse') occur in the DP ankles of primitive archosaurs. Development of a peg-and-socket articulation is a relatively minor innovation: it is certainly easy to envisage, and could have occurred in only one of two ways ('normal' or 'reverse'). Such minor features might well have appeared independently in distantly related archosaurs. Nonetheless these features are accorded primary significance in Chatterjee's analysis of archosaur phylogeny; by contrast some much greater and more complicated morphological changes (from

MT joint to CT joint and vice versa) are relegated to secondary importance.

By attaching primary significance to the course of the 'hinge' through the ankle (that is, to DP, MT and CT joints), it is possible to construct a more parsimonious scheme of archosaur phylogeny (Fig. 2*b*). In this scheme MT and CT joints are not derived from one another, but are derived independently from the DP joints of primitive archosaurs. Consequently there are no reversals in ankle structure, and convergences appear only in relatively minor features (the 'normal' and 'reverse' articulations of the proximal tarsals).

RICHARD A. THULBORN

Department of Zoology,
University of Queensland,
St Lucia, Queensland 4067,
Australia

1. Chatterjee, S. *Nature* **295**, 317-320 (1982).
2. Chaz, A. J. in *Studies in Vertebrate Evolution* (eds Joysey, K. A. & Kemp, T. S.) 121-155 (Oliver & Boyd, Edinburgh, 1972).
3. Thulborn, R. A. *Alcheringa* **4**, 241-261 (1980).
4. Krebs, B. *Paläont. Z.* **37**, 88-95 (1963).
5. Krebs, B. *Naturwissenschaften* **61**, 17-24 (1974).
6. Cruickshank, A. R. I. *S. Afr. J. Sci.* **75**, 168-178 (1979).

CHATTERJEE REPLIES—Thulborn wants to replace the term 'primitive mesotarsal' (PM) ankle joint¹ with 'duplex' (DP) ankle joint. I prefer the former term because the ankle joint in these groups (proterosuchians) is both functionally mesotarsal^{2,3}, and primitive as 'the astragalus and calcaneum articulate with each other by means of double convex-concave complementary joints'. Both of us agree that the PM (= DP) joint could give rise to the crocoid (CN and CR) joint. In his earlier paper⁴, Thulborn did not refute the argument that the AM (= MT) joint, as seen in dinosaurs, could have evolved from the crocoid joint, but he now considers that the AM joint was derived directly from a PM joint, not via a crocoid joint. In Fig. 1 (ref. 1), I suggested this latter possibility. But there are some dinosaurs which retain the vestiges of peg-and-socket joint, either in CN or CR fashion, indicating a close structural affinity. The origin of dinosaurs from different thecodontian lineages is not clear at this moment, but differences among the tarsi may yet be an important clue to such relationships. It is possible that carnosaurs evolved from two separate stocks of thecodontians (allosaurs from ornithosuchids, and tyrannosaurs from poposaurids), as indicated by the correlation between ankle structure and other anatomical details. In ornithosuchids and poposaurids, the initiation of the ascending process of the

astragalus, so prominent in later carnosaurs, can be seen (Fig. 1)¹. This also implies that the AM joint could have evolved from a crocodiloid joint.

SANKAR CHATTERJEE

The Museum,
Texas Tech University,
Lubbock, Texas 79409, USA

1. Chatterjee, S. *Nature* **295**, 317–320 (1982).
2. Cruickshank, A. R. I. *S. Afr. J. Sci.* **75**, 168–178 (1979).
3. Hughes, B. J. *zool. Res.* **156**, 457–481 (1968).
4. Thulborn, R. A. *Alcheringa* **4**, 241–261 (1980).

Reduction in plasma calcium during exercise in man

RUBEN and Bennett¹ reported increased plasma calcium concentration $[Ca_p]$ following exercise in various species. They suggested that increased $[Ca_p]$ is due to bone resorption, implying an increase in total plasma calcium content (Ca_t). However, the rise in $[Ca_p]$ following exercise might merely result from a reduction

in plasma volume (PV)² with or without change in Ca_t .

We have now investigated the responses of PV, $[Ca_p]$ and Ca_t in 13 men (22 ± 1 yr, 73.9 ± 2.1 kg and mean maximal oxygen uptake of 56.4 ± 2.2 ml kg⁻¹ min⁻¹) who worked on a cycle ergometer at one of three randomly ordered work intensities for 6 min on three different days (Table 1). PV was measured with Evans blue dye², and change in PV was calculated from the haematocrit³. $[Ca_p]$ was analysed by atomic absorption spectrophotometry. Ca_t was calculated from $[Ca_p]$ and per cent change in PV⁴.

These results, from men, show that during exercise, $[Ca_p]$ increased to a constant level while PV and Ca_t decreased progressively with increasing exercise level (Table 1). The rate of decrease in PV was greater than that for Ca_t (Fig. 1). This suggests that the increase in $[Ca_p]$ during exercise can be attributed to a more rapid loss of fluid from the vascular space and does not necessarily require a significant net influx of calcium from extravascular

Table 1 Effect of submaximal and maximal work on plasma volume and calcium ion concentration in humans (from Fig. 5 in ref. 2)

% Work load	% Net change in plasma volume	Ca ²⁺ concentration (mmol l ⁻¹)
Rest	—	1.10
43	-8	1.03*
62	-14	1.02*
100	-16	1.20*

* After 10 min exercise.

systematic lactic acid accumulation and pH depression associated with maximal levels of exercise in vertebrates. The plasma volume and calcium concentrations cited by Convertino *et al.* associated with 'heavy' muscular activity (225 W) represent submaximal exercise levels (about 60–65% of capacity) for humans. Thus, it is unlikely that exercise-related hypercalcaemia of the sort we described occurred in their experimental subjects. Significantly, Greenleaf *et al.*² have previously described a dramatic rise in the concentration of calcium ions in plasma immediately following maximal exercise, whereas there is a slight decrease in plasma calcium ion concentration at 43% and 62% work levels (Table 1).

Additionally, haemoconcentration cannot possibly account for the exercise-related hypercalcaemia measured in other species in our study. For haemoconcentration to account for the degree of post-exercise hypercalcaemia present in all non-mammalian osseous species, plasma volume reduction of at least 30% (for fish with acellular bone) and 50% (for fish with cellular bone and reptiles) would have to have occurred. Acute plasma volume losses approaching 20–25% are generally associated with hypovolemic shock. There were no indications that any of our experimental animals experienced such a condition. Moreover, the lack of exercise-related hypercalcaemia in non-osseous species investigated (lamprey, shark) seems particularly noteworthy in this respect.

While haemoconcentration may well be a contributing factor to some of the exercise-related hypercalcaemia we describe, it seems unlikely to be a major factor accounting for the phenomena discussed in our paper.

JOHN A. RUBEN

Zoology Department,
Oregon State University,
Corvallis, Oregon 97331, USA

ALBERT F. BENNETT

Department of Cell and Developmental Biology,
University of California,
Irvine, California 92664, USA

Table 1 Effect of light, moderate and heavy muscular activity on PV, $[Ca_p]$ and (Ca_t)

Exercise level (W)	PV (ml)	$[Ca_p]$ (mmol l ⁻¹)	Ca_t (mg)
Rest	3,504 ± 83	2.47 ± 0.03	347 ± 9
Light (100)	3,388 ± 91*	2.53 ± 0.03*	344 ± 9
Moderate (175)	3,208 ± 94*	2.55 ± 0.03*	324 ± 10*
Heavy (225)	3,072 ± 83*	2.57 ± 0.03*	317 ± 9*

Values are mean ± s.e.

* $P < 0.05$ compared with corresponding rest value.

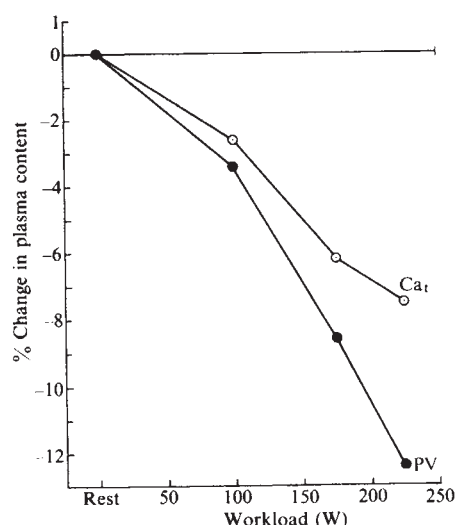


Fig. 1 Proportional changes in PV and Ca_t with graded work intensities of 100, 175 and 225 W. Plotted points are mean values of the percentage change calculated from pre- and post-exercise haematocrit and plasma calcium concentrations⁴.

sources. Thus, the conclusion of Ruben and Bennett¹ should be evaluated cautiously until measurements of vascular fluid and calcium shifts are available.

VICTOR A. CONVERTINO*

EMILY R. MOREY

JOHN E. GREENLEAF

Biomedical Research Division,
NASA Ames Research Center,
Moffett Field, California 94035, USA

1. Ruben, J. A. & Bennett, A. F. *Nature* **291**, 411–413 (1981).
2. Greenleaf, J. E., Convertino, V. A. & Mangseth, G. R. *J. appl. Physiol.* **47**, 1031–1038 (1979).
3. Van Beaumont, W., Greenleaf, J. E. & Juhos, L. *J. appl. Physiol.* **33**, 55–61 (1972).
4. Van Beaumont, W., Strand, J. C., Petrofsky, J. S., Hipskind, S. G. & Greenleaf, J. E. *J. appl. Physiol.* **34**, 102–106 (1973).

RUBEN and BENNETT REPLY—The values cited above by Convertino *et al.* for blood calcium levels during human exercise are not directly relevant to our previous study¹. We hypothesize that the mechanism resulting in post-exercise hypercalcaemia involves dissolution of a fraction of the crystalline calcium hydroxyapatite compartment of the skeleton. This dissolution is a result of

* Present address: Exercise and Sport Sciences Laboratory, Department of Physical Education, University of Arizona, Tucson, Arizona 85721, USA.

1. Ruben, J. A. & Bennett, A. F. *Nature* **291**, 411–413 (1981).
2. Greenleaf, J. E. *et al.* *J. appl. Physiol.* **43**, 1026–1032 (1977).

BOOK REVIEWS

Risk and rationality

Stephen Cotgrove

THE debate about risk can generate strong emotions — on both sides. Opponents of nuclear power, for example, are sometimes accused of being short on facts; or worse, they are charged with being emotional and irrational. Indeed, if we look at the risks of dying from various causes, from snake-bite to smoking, there doesn't seem to be much connection between probability and anxiety. One of the most persuasive arguments is that what we need is a calculus of risk — a league table — which will tell us when we should worry, and when to relax.

Such an approach, however, is simplistic in the extreme. Indeed, the two closely written books reviewed here are testament enough to the complexities of deciding on courses of action which have risky consequences. And at the end of the day, it is not surprising that neither comes up with even relatively simple solutions.

Of the two, *Acceptable Risk* is by far the more extensive in its coverage. As its title implies it analyses the processes by which we can arrive at decisions as to what risks are acceptable. Three main strategies are explored. First, there is formal analysis, of which cost-benefit analysis is the most familiar example. Second, there are what the authors call "bootstrapping approaches". In essence, these take the safety levels achieved with old risks as a guide to managing new risks, an example being the natural-standards approach which includes taking natural levels of radiation as the bench mark for judging acceptable levels. Finally, there is reliance on the theoretical judgement of the technical experts who are most knowledgeable in a particular field.

Each of these approaches has advantages and disadvantages, depending on the area of application. In deciding on the adoption of a new technology, for example, the formal approach is more open to evaluation and more politically acceptable than professional judgement. But this latter method is superior for arriving at routine individual decisions. The natural-standards approach assumes that past levels are currently acceptable. Inherent in each of these strategies is a number of generic problems which can be tackled with varying degrees of success. These include uncertainties in defining the problem, and difficulties in assessing the facts, and the relevant values.

Defining the problem includes identifying what consequences are to be taken into account. So, for example, the environmental movement has sought to

Acceptable Risk. By Baruch Fischhoff *et al.* Pp.185. ISBN 0-521-24164-2. (Cambridge University Press: 1982.) £15, \$19.95.
Risk/Benefit Analysis. By Richard Wilson and Edmund Crouch. Pp.218. ISBN US 0-88410-667-5; ISBN UK 0-66-60514-8. (Ballinger/Harper & Row: 1982.) \$23.75, £16.95.

extend the consequences of the impact of new technologies to include consideration of ecological consequences (changes to the balance of ecosystems). Setting the agenda is part of the political process: the definition of the problem reflects the prevailing balance of power. The way in which a problem is defined can even determine the outcome of the decision process.

The view that rational decision-making is mainly a question of getting the facts right also runs into considerable difficulties when we look at the uncertainties surrounding establishing the facts in all but the most straightforward examples. The greatest problems arise where there are insufficient cases to arrive at any reliable assessment of probabilities, though the results may be catastrophic. Where data are inadequate, fault-tree and event-tree analysis may be used to construct a model of a complex system which provides a basis for predicting the results of any one or more of a number of possible failures. In practice, such analyses are prone to a number of faults including failure to consider human errors. The DC-10 failures, for example, were the result of not recognizing the way in which a technological system functions as a whole. The simultaneous breakdown of safety systems designed to be independent (common-mode failure) has also constituted a major unanticipated hazard in several reactor accidents. And these are but a few of the difficulties pin-pointed by the authors.

A brief summary cannot do justice to the depth and breadth of the analysis laid out in *Acceptable Risk*. For example, it is not simply the probable frequency of the hazard, but the characteristics of the hazard which are crucial for understanding responses. Catastrophic risks generate more anxiety than more probable hazards with less disastrous consequences. Similarly, involuntary risks are less acceptable than voluntary.

A crucial factor in any decision on the acceptability of risk is the purpose for which the risk is taken. Where the end or goal is highly valued, such as saving a life, greater risks may be justified than for some

trivial result. Values are at the centre of all decisions entailing risks; but they do not easily lend themselves to quantification. While this question, a central one, is explored in some detail by Fischhoff and his associates, it is practically ignored by Wilson and Crouch.

Of the three strategies, cost-benefit analysis is the least likely to ignore values, but it can only treat economic consequences or those that are quantifiable in monetary terms. Nowhere are the difficulties more obvious than in the attempts to put a monetary value on life itself. Such attempts come up against a fundamental philosophical objection which is not sufficiently explored by the authors. To reduce all values to dollars assumes that this is the ultimate criterion for human conduct. The impossibility of putting a monetary value on human life should be warning enough that there are higher order values which cannot be traded against money. Yet it is no answer to say that because values are elusive and changeable they should be omitted from rational, objective analysis. This is a simple recipe for the uncritical acceptance of the values of the decision takers.

Fischhoff and his associates end their book with a chapter of recommendations for improving acceptable-risk decision making and with one identifying the areas requiring further analysis and research. In all they have achieved a masterly systematization of the many practical, methodological and philosophical problems raised by attempts to answer the question "How safe is safe enough?"

In *Risk/Benefit Analysis*, Wilson and Crouch tackle a more limited brief. They too emphasize the problems inherent in defining, perceiving and estimating risk. They stress the importance of drawing the attention of the decision-makers to the uncertainties in the process and of emphasizing the distinction between decision-making and risk analysis. But the authors explore a more limited range of strategies and lack an awareness of the philosophical presuppositions, political acceptability and underlying assumptions about human behaviour which are such an important element in *Acceptable Risk*. The strength of their book is in its exposition of the practice of risk analysis, and in its actual review of nine case studies, most of which, however, fall far short of the rigorous and systematic approach set out in the text. This book, too, constantly underlines that risk analysis is only an aid to

decision making and not a substitute for it.

Part of the reason for the growing anxiety about risk is obvious enough — modern technology holds out great promise but it carries with it potential for danger. Its benefits and hazards are not equally distributed. Nor is exposure to risk always within the control of the individual, who may choose not to smoke, but has little say in the disposal of toxic waste.

The importance of both books is to emphasize that rational decisions on risk

cannot be reduced largely to a calculus of probabilities. Such an approach would simply brush under the carpet the inescapable political dimensions of many such decisions. But worse, it fails to account for the ways in which reasonable and rational beings in fact seek to cope with the complexities of how to act in the face of potential danger. □

S.F. Cotgrove is Professor of Sociology at the University of Bath.

Dendrochronology by numbers

John Fletcher

Tree-Ring Dating and Archaeology. By M.G.L. Baillie. Pp.274. ISBN UK 0-7099-0613-7; ISBN US 0-226-03630-8. (Croom Helm/University of Chicago Press: 1982.) £16.95, \$25.

THE attraction of using tree-rings for dating buildings and archaeological artefacts lies in the high accuracy — to the year or to within a decade — that can be achieved if the felling date of the timber can be identified. But how often is that feasible? One limitation lies in the relatively few species which have annual growth that can be reliably identified and converted to a chronology. Another is that in temperate areas various ecological factors affect the growth rings of each species, so a range of reference chronologies may be necessary to permit a large fraction of the somewhat different ring-width patterns to be matched. In addition, palaeoclimatologists recognize from climatic swings in the past that the extent of dendroecological zones has not remained constant.

Artefacts from extreme environments are especially suitable for tree-ring dating as stress comes mainly from one factor. Douglass in 1929 was the first to use this to advantage and date ancient cliff dwellings in Arizona and New Mexico from the pines used in building them. In central Europe, first Huber then Hollstein had by the mid-1960s matched sequences of rings from several hundreds of the oaks used in buildings and artefacts in hilly areas, including prehistoric lakeside settlements around the Alps.

In contrast, the diverse stresses which apply to valley and lowland oaks in north-west Europe made the pattern matching of rings impossible in the pre-computer era. In this book Dr Baillie tells of his success, using a computer program devised with his colleague J.R. Pilcher and now widely adopted, in matching Irish samples and a few from south-west Scotland. Archaeological remains of oak are scarce in Ireland, and the items dated number only about thirty. They include cruck houses, the Dublin waterfront, lake dwellings known

as crannogs and the archaeologically interesting horizontal water-mills. The results have enabled Baillie to construct two oak chronologies for the last two millennia (that for fossil timber for carbon-14 calibration goes back much further).

The methodology described is largely that of the Tucson tree-ring laboratory in which ring-widths are converted to indices, thus losing, when plotted, visual features known as indicator years and signatures. (As applied elsewhere in Europe these features are of particular value, for example in matching short sequences with dates known approximately by other evidence.) The author's approach is largely statistical and he displays no interest in the scientific and analytical research which has revealed relevant information about wood structure and tree physiology. Thus the value of a rare abnormality in oak, noticed by Buffon as of chronological significance, is not mentioned.

The book gives a detailed, step-by-step account of the regional application of the author's method to a limited number of widely separated items in the wetter zone which covers much of the western part of the British Isles. As such it will be of interest to departmental libraries and to the enquiring archaeologist. It is not suitable for use as a text-book, however. Some statements are misleading in the context of European dendrochronology, while inaccuracies render parts unsuitable for being quoted. This applies in particular when the author turns from Ireland to write briefly about southern Britain. In this region, not only oak trees but oak artefacts and excavated material are numerous. In the lower and drier zone — which includes rain shadows in the midlands, the east and south-east of England, as well as continental areas around the southern basin of the North Sea — the Irish chronologies rarely apply (and vice versa) while Baillie's method is too limited for success in matching a high proportion of samples. □

John Fletcher is at the Research Laboratory for Archaeology and the History of Art, University of Oxford.

Course of storms

T. N. Krishnamurti

Tropical Cyclones: Their Evolution, Structure and Effects. By Richard A. Anthes. Pp.208. ISBN 0-933876-8. (American Meteorological Society, Boston: 1982.) \$40.

IN TEACHING tropical meteorology courses it is generally necessary to devote about 25 per cent of the course to tropical cyclones: their origin, maintenance and eventual decay. Richard Anthes has provided, for my needs, an important text that covers this area more than adequately.

The monograph deals with the structure, life cycle, physical effects, numerical simulation, storm modification, oceanic responses and weather forecasting aspects of tropical cyclones. The description of structure is based on the best available observations of such storms, mostly provided by the National Hurricane Center in the United States. In the area of physical effects the role of cumulus convection is brought out in considerable depth, and a critical appraisal of the convective heating, the vertical eddy flux of heat, the role of non-convective heating, convergence of eddy flux of moisture and the moistening by cumulus scale motions is presented neatly and clearly. In this area, Anthes discusses a number of recent findings, most notably the role of radiative forcing in the diurnal modulation of tropical cyclones. His breakdown deals with observational as well as modelling aspects, and will be most helpful to students modelling such storms.

The account of modelling covers the state of the art of simulation of idealized as well as real data numerical weather prediction experiments in two and three dimensions. Several examples carry important information on why storms form from incipient disturbances in some cases and not in others; this is largely based on Anthes' own research over many years. The oceanic response is a topic hardly addressed in most other texts or reviews, and is a welcome addition to the completeness of this book. The noteworthy area here is the role of the oceanic stratification.

The final section covers the use of statistical and dynamical methods in operational prediction up to three days from the initial state. Here Anthes summarizes the current ranges for the error statistics in hurricane (typhoon) prediction centres in the United States and elsewhere.

I find this an extremely important book with few areas to dispute in the assessment of our current understanding of hurricanes. No previous text has addressed the wide range of teaching needs covering both observations and theory. It is a blend that is remarkably well done in this monograph.

T.N. Krishnamurti is a Professor in the Department of Meteorology at The Florida State University, Tallahassee.

Antarctica: the geological background to resource assessment

J.F. Lovering

Antarctic Geoscience. International Union of Geological Sciences: Series B, No.4. Edited by Campbell Craddock. Pp.1,172. ISBN 0-299-08410-8. (University of Wisconsin Press: 1982.) \$35, £26.25.

THE last continent — Antarctica — is rapidly assuming a new importance as a resource-hungry world begins to wonder whether it hosts exploitable mineral and fuel deposits and, if it does, who has title to them. Since 1961 the nations signatory to the Antarctic Treaty have quietly been engaged in basic geological and geophysical studies of the continent and its ice blanket. Although their research has not been directed specifically towards resource exploration, it is reasonable to conclude that many programmes have more than a passing relevance to an assessment of Antarctic resource potential.

For these reasons a lot of people, not all of them geoscientists, have been waiting with impatience for this massive volume which contains the papers presented at the Third Symposium on Antarctic Geology, together with a 1:5,000,000 geological map of Antarctica compiled from data available up to 1972. Their impatience is well justified since the symposium was held in Madison, Wisconsin, in August 1977.

In virtually any other scientific field, a symposium volume having a gestation period of five years would, on its birth, be hopelessly out of date. This volume is saved from such a fate by the studied pace with which field studies in Antarctica must proceed — the self-evident climatic constraints limit activity to around three months each austral summer so the rate of data collection is relatively slow. Still, it is difficult to understand why it has taken so long for the book to appear. The editor's preface hints at a multitude of problems and at his growing horror as he gradually recognized the enormity of his task. We should be grateful that he persisted to the bitter end.

Antarctica is the keystone to reconstructions of Gondwana so it is not surprising that many contributions focus on the specific relationships between Antarctica and South America, Africa, India, Australia and New Zealand. Many reconstructions use ocean floor magnetic lineation data which were relatively new at the time of the symposium but are now merely records of the state of the art at that time. Probably the most valuable series of papers in this general area is that concerned with the nature and evolution of the Scotia Arc — that puzzling and highly deformed connection between the margin cordillera in southern-most America and the Antarctic Peninsula. The islands of the Scotia Arc (South Georgia, South Orkney Islands and South Shetland Islands) are shown to have been initiated in Middle to

Late Jurassic times on an older metamorphosed "basement" complex, probably in an old re-entrant in Gondwana which formed a physical separation between the Andes and the West Antarctic cordillera. Other contributors, however, suggest that the eastern South Scotia ridge is an intraoceanic island arc.

In view of the recent fracas in the Falkland Islands, and the often-quoted petroleum potential of the sedimentary cover over the Precambrian basement of the area, many readers will be disappointed that only one contribution discusses Jurassic, Cretaceous and Tertiary sequences on the Falkland Plateau. Nonetheless they will be interested to learn of the existence of a sapropelic claystone with an organic content averaging three per cent within the Jurassic sediments near the base of the cover section. These same people will also be intrigued by the records of the first multichannel seismic reconnaissance survey of the thick sedimentary sequence filling the Weddell Sea basin, an area thought to have high petroleum potential on previously rather general grounds. There are enough promising structures and traps indicated on the records to send exploration managers hot foot to their friendly seismic survey contractor.

Another series of papers is concerned with the constitution, geochronology and structural evolution of the East Antarctic Precambrian Shield. These contributions represent a significant advance in our understanding of this important Archaean Shield, but still form only a relatively crude picture of what is a structurally complex but fascinatingly accessible example of a very ancient piece of the Earth's crust. Resource implications of the East Antarctic Shield rocks receive little attention except for one paper on the Precambrian "iron deposits", more realistically termed jaspillites, of the Prince Charles Mountains. With an average iron

oxide content of 43 per cent it will be a long time indeed before anyone looks to Antarctica as a source of supply for the world's steel makers.

Only two other papers are concerned with metallic mineral resources, and both examine the possible association of base metal deposits with the igneous rocks of the Antarctic Peninsula. Although many have suggested the existence of mineral wealth in the Peninsula by analogy with the mineral-rich central Andes, a warning is given that this conclusion may be premature because the two regions have rather different Cenozoic tectonic histories.

There is much that I have not discussed and it is simply impossible to do justice to the rich variety of information contained within this volume — including structural geology and tectonics, crustal structure, palaeontology, marine geology and Precambrian, Palaeozoic and Cenozoic history. I have focused on those studies which have a bearing on the future assessment of Antarctic resources both on- and off-shore, but the ice flow and bedrock information arising from the beautiful data provided by air-borne radio echo-sounding should at least be mentioned.

The confirmed Antarctic geoscientist will need this volume (and, at its remarkably low price, probably buy it); but the book hardly constitutes an easy entry to the subject for the non-specialist. I would like to see a wider group become interested in the geological evolution of the continent since Antarctic geologists and geophysicists comprise a small group largely closed to outsiders who have not braved the rigours of an Antarctic field season. We still badly need an up-to-date and readable monograph on the subject to stimulate this wider interest. □

J.F. Lovering is Professor of Geology and Chairman of the School of Earth Sciences at the University of Melbourne.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Wealth beneath the ice? Heavily glaciated terrain in the Transantarctic Mountains, dividing the East and West Antarctic plates and bordering the Ross Ice Shelf.

David Drewry, Scott Polar Research Institute.

Semiconductors from Bell Laboratories

Andrew Holmes-Siedle

Physics of Semiconductor Devices, 2nd Edn. By S.M. Sze. Pp.868. Hbk ISBN 0-471-05661-8; pbk ISBN 0-471-05983-7. (Wiley: 1982.) Hbk £37.20, \$63.20; pbk £12.91, \$23.30.

THE impending publication of a new edition of a famous textbook is always an occasion for keen anticipation. In the field of semiconductor devices, there have been few textbooks more authoritative or comprehensive than the first edition of *Physics of Semiconductor Devices*, a book which bore the stamp of the author's broad personal experience of device development and the unique aura of Bell Laboratories. That 1969 edition, however, was clearly becoming dated by the rapid advance of integrated semiconductor device technology. The performance and sizes of devices had altered by orders of magnitude, while the principles of device operation had been the subject of over 40,000 research papers. Obviously, the revision of so comprehensive a book would be a huge undertaking.

I can only say that the task has been completed with a thoroughness that leaves me open-mouthed with admiration. All the parts of the first edition which I had earmarked as obsolete have been eradicated, and the structure has been subtly altered to accommodate the changes of the past 15 years. Dr Sze states that over 65 per cent of the references and diagrams are new and 80 per cent of the text has been revised. All this has been done most effectively; the mark of authority again appears in the large amount of new material added.

The first edition was divided into five major sections: Physics; p-n Junction Devices; Interface Devices; Optoelectronic Devices; and Bulk-Effect Devices. In the second edition, the major growth of metal-oxide-semiconductor device technology has been acknowledged by means of a new division of sections between bipolar and unipolar devices. Microwave devices also have their own section, while photo-detectors and solar cells are now rightly given chapters to themselves. The broadening of the treatment of solar cells will prove extremely useful to many students of energy generation: it goes so far as to give some useful data on geographical distribution of solar flux and its spectral properties, relates these to mechanisms of conversion in semiconductors and adds a fascinating account of the new device structures being investigated (for example the textured single-crystal surface, the cascade two-junction cell, thin films etc.).

In the section on metal-oxide-semiconductor devices, the author evidently felt that there was a danger that the immense body of recent information could run away with the book; for instance, there is a notable austerity in the handling of MOS measurement methods. In the section on the myriad forms of the MOS field-effect transistor (the device which has revolutionized computers, watches, entertainment and communications), the new relevance of the submicron device structure is well described. Surprisingly, the author manages to confine the growth of these sections to an increase of about ten pages. At the same time little that is really important has been left out, although (inevitably) details of processing and circuitry have had to be ignored. Real flaws

are rare. The brief treatment of ionizing-radiation effects is not up-to-date; so much has happened since the 1967 references given. Also, it is probably wrong to say that carrier generation in SiO_2 is caused "by breaking Si-O bonds".

While such a concentrated treatment of semiconductors could never be light reading, Dr Sze has certainly taken pains to present his material clearly, with liberal use of line drawings as explanations in their own right. As a result, the book will remain an indispensable companion to the student and research worker. Moreover, I believe it will be illuminating for journalists and others who have to know what this commercially-important field of semiconductors is about. □

Andrew Holmes-Siedle is Consultant in Solid State Physics and Radiation Damage to the Fulmer Research Institute, Stoke Poges, Buckinghamshire.

Fingers on the pulse of the insect

John Brady

Insect Clocks, 2nd Edn. By D.S. Saunders. Pp.409. Hbk ISBN 0-08-028848-0; pbk ISBN 0-08-028847-2. (Pergamon: 1982.) Hbk £45, \$90; pbk £24, \$48.

FEW scientific monographs survive to a second edition, and still fewer get there within six years of the first. It is a measure of the success of *Insect Clocks* as a textbook, and of the general interest in this apparently esoteric corner of biology, that the book has already been given a face-lift.

The interest in insect timekeeping stems from insects' economic importance and the tantalizing hope of distressing their life cycles in our favour by manipulating their clocks. It must be said, however, that basic knowledge of their timekeeping — as evinced by the 1,120 papers Saunders reviews — far outstrips any practical manipulation thereof. This is in marked contrast with the commercial application of false photoperiods to induce plants to flower or poultry to ovulate year in and year out. The unfortunate truth is that 99 per cent of insect pests occur in the open air.

If justification for this second edition were needed it is that since 1976 a number of gaps in our knowledge have been closed. An important one is the apparent reconciliation of "hour-glass" and "circadian" explanations of insect photoperiodic control. The former finds daylength apparently being measured by a single, non-repeated process (like a kitchen timer), the latter that organisms seem sensitive to light at a particular point in their daily cycle, every day (like a 24-hour light-sensitive clock).

One should never expect simple answers

in biology, however, and having argued the reconciliation in depth, Saunders is then forced to largely demolish his case with an added footnote saying that A.D. Lees's latest evidence from his aphids can still best be described as an "hourglass".

That much less is known of the ultimate chemistry and physiology of insect clocks than of those of micro-organisms or molluscs is as apparent in this second edition as it was in the first. Most research over the past six years has been along the same lines as before, dealing with phenomenology and modelling of insect timekeeping, and it is with this that the book's 400 new references are chiefly concerned. Nevertheless the — albeit modest — advances in knowledge of the clock's coupling mechanisms do seem underplayed. The greatest single addition to the book is a long new chapter on multioscillators.

All is admirably explained in Saunders's usual lucid style, and this time he has removed some of the unnecessary algebraic symbols that speckled the first edition. The book does still suffer slightly from its original division into chapters on rhythms in single insects, rhythms in populations and rhythms in physiology, with the result that entrainment, phase response curves and so on have to be dealt with repetitively rather than as the single phenomena they are. *Insect Clocks* is, however, well on the way to becoming a classic, and this new edition should surely find its way onto the shelves of insect physiologists and chronobiologists alike. □

John Brady is Reader in Insect Behaviour in the Department of Pure and Applied Biology, Imperial College, University of London.

● Volume 3 of *Electron Microscopy of Proteins*, edited by James Harris, has recently been published by Academic Press. Price is £27.80, \$57. For review of Vols 1 and 2 see *Nature* 298, 498; 1982.

NEW WAYS WITH MAGNETS

The case for attending to magnets

LIKE the air we breathe, magnets and the materials of which they are made have become ubiquitous but almost unobtrusive. The pages that follow are not meant as a systematic survey of what has happened since the first (probably Mediterranean) navigators used naturally occurring magnets in steering their ships, but as a pointer to the interest of this field in the past few years. For it tends to be forgotten how much things have changed in the years that have followed the Second World War.

Part of the difficulty is that magnetism itself retains some of the mysteriousness that perplexed the old mariners (or should have done). Magnetic forces are non-central. A sufficiently small dipole (say a nuclear spin or an electronic magnetic moment) will neither attract nor repel another like it, but create a torque. And it is not even now surprising that the force on a charged particle travelling in a uniform magnetic field should be perpendicular both to the field and to its own direction of motion? Faraday's description of this unexpected behaviour 150 years ago remains a remarkable exercise in thinking the unthinkable.

In the circumstances, it is not surprising that the most obvious properties of magnetic materials, the ferromagnetism of iron and nickel, for example, should have been explained (and only then phenomenologically) only when an account of the electronic structure of atoms had been completed in the second half of the 1920s. And although it is now natural to look askance at atoms of transition elements in which outer electronic shells have begun to be filled while inner shells are still incomplete, the calculation of the most familiar properties of ferromagnetic materials — say the Curie temperature above which permanent magnetism disappears — remains an empirical art. The best that can be done is to show by model calculations that such a transition is an order-disorder transition or, more accurately, that model lattices such as the Ising lattice can produce behaviour of the required kind. (That the Ising lattice, a set of interacting bi-directional spins at the corners of a two-dimensional square lattice, can be solved exactly has been a powerful but frustrating challenge for those who would solve real models, but has engendered a whole new industry for theoretical physicists.) The consequence, however, is that the search for new magnetic materials has become an empirical process guided only by

enlightened guesswork (see box).

Even so, there have been some important successes. Néel's work on the ferrites twenty years ago is one vivid proof of how phenomenologically-based theory can help to illuminate even problems not exactly calculable. Much the same is true of the way in which the study of magnetic domains, and of the migration of their boundaries when the magnetization of the material that contains them changes, has pointed to the design of materials now almost certain to play an important part in emerging technology.

Exactly the same circumstances obtain in the development of new superconducting materials. That it is even now known that some materials (tin, niobium and so on) literally lose electrical resistance at sufficiently low temperatures is something to be grateful for. But while the essential character of the phenomenon has been understood for the past twenty years, the prediction of the properties as a superconductor of some new alloy remains largely a matter for empirically guided speculation. The underlying difficulty is the complexity of the real world.

None of this has prevented the steady march of magnets in the technology of research and in public use. The dependence of experimental high-energy physics on the improved design of magnets (see page 665) is unsurprising, but it tends to be forgotten that analytical techniques such as mass spectrometry and nuclear magnetic resonance are no better than the magnets around which the instruments are built. Fortunately, in this respect at least, empiricism has almost entirely given way to computation as computers large enough to handle real-life models have arrived.

With all this said, it is surprising that the whole field of magnetism remains as neglected as it is. While the company of those with an interest in designing better magnets, or understanding how existing magnets work, is quite substantial, their preoccupations are not generally shared. The unfamiliarity of the subject is as forbidding as its intrinsic difficulty. (For experimentalists, mere measurement remains a seemingly needless bugbear.)

Yet the way in which the technological applications of magnets in the past few decades have improved the performance of saleable devices, and made others feasible, suggests that the time has come when the research community and its sponsors should take the subject more seriously. An element of science policy in search of a government, perhaps? □

Materials galore

LODESTONE, the archetypal permanent magnet, was probably another name for magnetite (Fe_3O_4), which curiously has been the basis for most recent excitement about the development of magnetic materials. For magnetite has proved to be a literal representation of what the old school textbooks of chemistry say — a molecular compound of two iron oxides (FeO and Fe_2O_3) in which divalent and trivalent iron atoms occupy different kinds of lattice sites. The ferrites are merely lattices like those in which some divalent iron ions are replaced by others with different electronic properties.

The consequences of ferrite revolution are apparent inside any transistor radio. The old metallic magnetic components, in transformer cores for example, have been replaced by cylinders of dullish ceramic material which have in turn been fabricated from raw oxides or carbonates of their component elements. But the same ferrite revolution has helped also to throw light on what makes a material ferro or ferri magnetic, as well as providing Professor Louis Néel with a Nobel for his part in the explanation.

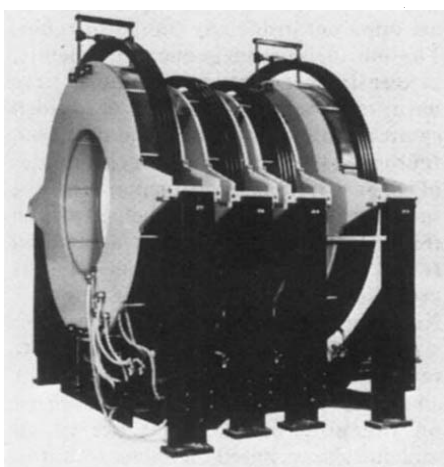
So what can be next? With the ferrite revolution now reduced to the search for small improvements, attention is focused on the potential of what are called metallic glasses based on oxides of boron and phosphorus heavily doped with oxides of iron and nickel. Being amorphous, these materials are free from the hysteresis that arises in crystalline materials because magnetic domains can be smaller than single crystals. The metallic glasses may prove to be easily and cheaply manufactured, leading to even less "lossy" power transformers and magnetic tapes that can be read more faithfully than those now on the market.

The magnetic tape industry seems to provide the chief incentive for another current development — that of alloys of cobalt and chromium with high coercivity. The technical objective is to allow some form of recording information by means of magnetization perpendicular to the plane of the recording tape, and is made possible by the high coercivity (resistance to demagnetization by magnetic fields) of these alloys. The result should be recording tapes with improved sensitivity, a denser packing of information and a more nearly digital representation of it.

Superconducting magnets

No longer are superconducting magnets confined to research laboratories or those which use nuclear magnetic resonance (NMR) machines for more routine measurements. With the development of techniques that, for example, permit the use of NMR in the study of living tissues, it is natural that the body-scanning fraternity should have devised machines intended to provide biochemically based whole-body scans of human beings. This development entails the construction of magnets providing a uniform magnetic field of 2 tesla or so over regions 100 cm in dimension — for practical purposes unattainable without the use of superconducting magnets.

One of the British companies active in the field, Oxford Instruments Limited, has already built a number of magnets of this type. Physiological whole-body scanning is for the time being directed at two objectives — the use of the resonances of ^{31}P as indicators of physiological function,



A 0.15-tesla four-coil resistive magnet for whole-body imaging. Clear-bore access is 800 mm.

and of proton resonances as indicators of where water is more plentiful in the body (which could be an alternative to the use of diagnostic X rays).

Whether large-volume magnetic fields for such purposes will prove to be in great demand will no doubt hang on the outcome of the several clinical studies with prototype machines now being planned. The magnet manufacturers will be overjoyed if it turns out that scanning procedures yielding representations of the structure of cross-sections through a human body (such as may be constructed from suitable proton magnetic resonance measurements) are preferable to the more familiar use of diagnostic X rays. Meanwhile, the supply of large-volume magnets for the prototype machines, at a cost of £100,000 or so each, seems to be a sufficient recompense for the high cost of development.

In the evolution of superconducting magnets for laboratory use, the cutting edge of development seems to be what it

has been for the past twenty years — the improvement of magnets for use in nuclear magnetic resonance. For the higher the value of the magnetic field in which a sample is immersed, the greater the resolution that will ultimately be possible. But the sharpness of the spectral lines that eventually emerge will also be determined by the uniformity of the field over the volume of whatever sample is being used, with the result that stronger fields must also be more uniform.

Standard superconducting magnets with pre-specified field homogeneity can now be bought off the shelf. Central field strengths of these standard products range up to 14 tesla (140,000 gauss), but the magnets themselves may be no more than a foot long. The compactness of the range of superconducting magnets now available is perhaps their most striking characteristic, especially when the imagination is invited to speculate about the bulk of the power supply equipment that would be necessary to generate fields of a tesla (the maximum obtainable with conventional electromagnets).

Cooling problems

The analogous penalty in the use of laboratory superconducting magnetics is, however, the cryostatic equipment needed to maintain the superconductor at liquid helium temperatures. The coils of a superconducting magnet must be immersed in liquid helium or, alternatively, in liquid-helium vapour. Naturally enough, users of superconducting magnets have no wish to double as low-temperature physicists, worrying about the adequacy of the cooling system. Fortunately, most of the suppliers of laboratory magnets have been able to reduce the rate at which helium is lost from their cryostats to the point at which topping up the system with liquid helium is necessary only at intervals of a few months. The ideal is that replenishment can be tied in with routine maintenance visits.

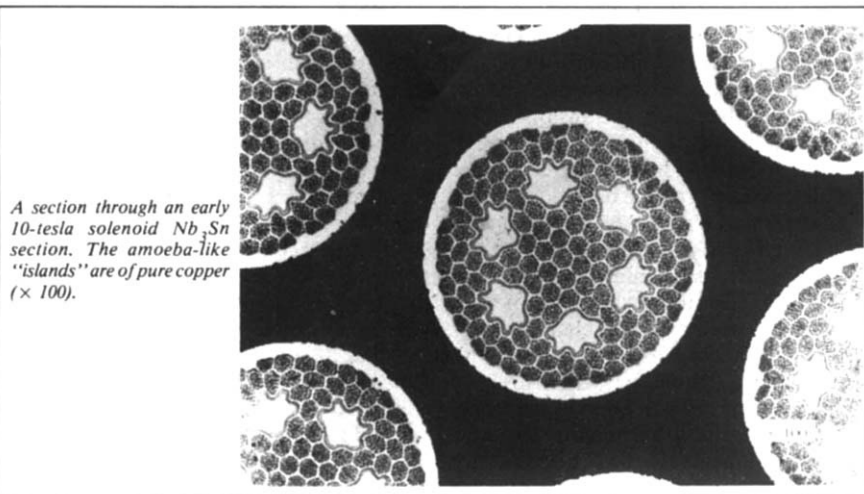
For the rest, the design of a super-

conducting magnet raises few problems not encountered by the traditional designers of electromagnets. The simplest magnets are solenoids. Often it is convenient to split the solenoids in half, on a plane perpendicular to the axis, but at the expense of maximum field strength. Improving the homogeneity of the field, increasing the volume over which irregularities do not exceed some specified percentage, is to some extent a function of the way the coils are wound, but the inclusion of shims of superconducting material is usually also advantageous.

Quite apart from considerations of compactness, the overriding attraction of superconducting magnets over conventional electromagnets is that the maximum field is not limited by the magnetic permeability of the core material, soft iron perhaps. Yet there are limits to the extent to which the field strength of a superconducting magnet can be increased which are themselves determined by the properties of the superconductors used for constructing the coils. The crucial consideration is that the material of which the coils are made should not experience magnetic fields great enough to rob them of the superconducting properties. With existing superconducting alloys (niobium-titanium and niobium-tin) it should be possible to push beyond 17 tesla, the present limit of what is practicable.

The most straightforward way of using this material is to deposit it on layers of a tape of copper and stainless steel (relatively an insulator at liquid helium temperatures, but a means of carrying current if superconducting properties should at any point be lost). More recently, techniques have been developed for forming niobium-tin alloy from filaments of niobium which are annealed with bronze in such a way that the tin component of the bronze preferentially alloys with the niobium. The residual copper serves both as a mechanical support and as a means of conduction should superconducting properties be lost.

Where the design of superconducting magnets will go from here is anybody's guess. The few companies specializing in laboratory-scale magnets (beam-forming



A section through an early 10-tesla solenoid Nb_3Sn section. The amoeba-like "islands" are of pure copper ($\times 100$).

magnets are another business) usually offer custom-built designs intended to suit the needs of people in academic or industrial laboratories in need of some predetermined configuration of magnetic field. As with NMR, from any one of these applications might spring some new range of off-the-shelf superconducting magnets.

Oxford Instruments, one of the principal suppliers, says that it will consider *ad hoc* proposals from potential academic customers, but will not in the first instance produce a detailed estimate of the cost of building a new design. Rather, it will undertake to produce information sufficient to sustain a research grant application, and will produce a final estimate of cost in return for a fee refundable against a firm order when the application has been successful.

Whether this esoteric business will — or would be — transformed by the development of superconducting materials with transition temperatures substantially above the boiling point of liquid helium is another matter. The chief benefit, as things are, would be quantitative: helium consumption would be reduced. But if somebody were to construct a material superconducting at liquid nitrogen temperatures, a whole new ballpark would be accessible. . . . □

Cold comfort at room temperature

SUPERCONDUCTIVITY is the next best thing to perpetual motion: an electrical current once established in a superconducting circuit will keep going indefinitely. So is it not mere misfortune that the only circumstances in which known materials can be made superconducting are those involving temperatures in the liquid helium range, below 4.1 K or thereabouts?

After the best part of a decade, the hopes first fostered by the empirical discovery that suitably constructed linear polymer materials have electronic conduction bands and are indeed capable of conducting electricity might make possible the design of molecules exhibiting superconductivity at room temperatures. In reality, painful experience has shown that superconductivity is doggedly a low-temperature phenomenon, involving lattice or molecular vibrations (phonons) and electron movement in a cooperative transition from a disordered to an ordered state. Room-temperature superconductivity is likely always to lie in the future.

For practical purposes, it is a sufficient but still daunting goal to find materials

superconducting at the temperature of liquid nitrogen, the common outer refrigerant in superconducting cooling systems.

The record so far is a transition temperature of 21 K for a version of Nb₃Ge in which germanium atoms are partly replaced by atoms of aluminium. (This is a true superconductor of Type II, able to sustain high magnetic fields without losing its essential properties, but metallurgically hard to work with.) As things are, such interest in the search for improved superconductors centres on the sulphides of molybdenum doped with rare-earth elements as substitutes for Mb.

In practice, even unexpectedly rapid progress in the search for superconductors with higher transition temperatures would only slowly transform the technology of making superconducting magnets. In practice the improved superconductors now coming into use are metallurgically almost impossible. They function as planned only if pure and stoichiometrically exact. It will be some time before it is a matter of routine to wind kilometres of wire on some form and call the result a magnet.

High-energy physics rides on flux

THERE is a real sense in which the development of high-energy physics in the past half century has been sustained by the technology of the magnet builders. Although the first important landmark in the artificial manipulation of nuclear structure — the use of 300 MeV protons for the disintegration of nuclei by Cockcroft and Walton in 1931 — was made possible by means of an electrostatic accelerator, E.O. Lawrence was even then building the first cyclotron at Berkeley. By the time, in 1935 or so, that the limits of electrostatic acceleration had been reached, working cyclotrons were all the rage.

The principle of these machines is simple enough. The path of a charged particle travelling in a magnetic field will be curved, but no kinetic energy will be lost. If the magnetic field is sufficiently extensive, the particle will travel in a closed path, thus increasing the length of time for which other means of accelerating it can do their work.

The cyclotrons which in the 1930s followed the Berkeley design were the simplest of all accelerating machines. The principle is simply to construct a substantial volume of uniform magnetic field between the poles of an electromagnet, to provide a source of charged particles at the centre of the arrangement and to arrange for some means of adding to the kinetic energy of the charged particles as they traverse closed orbits between the poles of the source magnet.

Cyclotrons have functioned since the

early 1930s because the frequency with which a charged particle traverses its closed orbit in a magnetic field is independent of the dimensions of the orbit and simply a function of the characteristics of the particle and the strength of the field (B). Thus the frequency of rotation is given by the familiar relationship $eB/2\pi m$, where e and m are the charge and mass of the particles.

Since the beginning, cyclotrons have been no better and no worse than the design of their magnets. Since the first machine at Berkeley, the standard procedure has been to accelerate charged particles by means of alternating electrostatic fields within a pair of semicircular hollow conducting surfaces shaped so that, when put together, they would constitute a hollow pancake-shaped structure — and which, when separate, are known as “dees”. The two halves of this structure are fed with radiofrequency power at the characteristic frequency of the cyclotron (a function only of B and the particle concerned).

One obvious difficulty with which the magnet designers have had since the beginning to contend is the sheer physical problem of shaping the magnetic field appropriately. Since charged particles spiral outwards from the central point at which they are injected, the field must be for practical purposes uniform over the maximal orbits followed by the particles. The construction of electromagnets spanning close on 2 m (such as with the 184-inch cyclotron built at Berkeley after

the Second World War) is a formidable undertaking.

For one thing, the sheer bulk of the magnetic circuit required to provide a return path for the magnetic field between the poles of the electromagnet where the action is is almost a feat of mechanical engineering in its own right. For another, the shaping of the pole pieces (usually machined from cast or forged blocks of soft iron) is a complicated process of balancing the loss of flux through the sides of tapering circular endpieces against the need for uniformity of field.

Throughout the 1930s, the construction of cyclotron magnets was almost a black art. Once a machine had been built, the designers of the magnet would be required to fiddle around with the placing of shims near the poles of the magnets to improve the characteristics of the field, its uniformity in particular. Even the first cyclotrons, however, were blessed with one important built-in benefit — because, between the poles of an electromagnet, flux lines unavoidably become more distant from each other, displacements of circulating charged particles towards one pole or another are discouraged by the mirroring effect thereby produced.

For practical purposes, cyclotrons have been used for accelerating protons and heavier ions, with maximum energy (determined by the area of the space between the pole pieces) measured in tens of MeV. The acceleration of electrons to comparable and greater energies is best accomplished by means of linear accelerators (most spectacularly by the two-mile accelerator at Stanford,

California) but the brief wartime history of the machine called the betatron is a landmark in the ingenious development of accelerating magnets.

The principle of this machine is that particles would be made to travel in strictly circular (and not outward spiralling) orbits, and would be energized by means of a varying magnetic field. In principle, a betatron is a transformer in which the primary winding is the wiring that energizes the electromagnet and the secondary "winding" is the circulating beam of electrons. Betatrons thus exploit the principle of magnetic induction, providing an electric field tangential to the circular orbits of the particles which is itself proportional to the rate of change of the magnetic field.

The evolution of this machine was largely the achievement of D.W. Kerst, who built a 2.3 MeV electron accelerator at the University of Illinois in 1940, who afterwards worked at General Electric on the development of increasingly energetic electron accelerators for use in medicine and other applications and who returned to the University of Illinois to complete (in 1961) a betatron with a maximum energy of 300 MeV. Inevitably, the operation of such machines is limited by the rate of energy loss by radiation from circulating electrons. Technically, the magnet design (at least as practised by Kerst), involving the use of laminated steel to avoid energy loss by means of eddy currents, seems to have been relatively predictable.

The development of high-energy physics has since then been determined entirely by the design of accelerating machines in which particles travel in or near predetermined orbits. The advantages are principally that the volume within which it is necessary to provide a magnetic field is confined to an area small compared with the area covered by the orbit of the accelerating particles. But the magnetic field must be steadily increased as the momentum of the particles is increased, by means of electric fields generated in oscillating radio-frequency resonant cavities for example. The result is a generation of machines called synchrotrons which produce a pulse of

energetic particles after each cycle of pre-programmed increase of the magnetic field around the orbit.

For practical purposes, electron machines such as that at DESY in Hamburg are simpler than those designed for accelerating protons, for provided that the source of particles to be accelerated is already relativistic, the frequency of the source of accelerating energy can be virtually constant. Proton synchrotrons must be designed in such a way that the frequency source has a dynamic range of frequency at least great enough to span the total range of acceleration of the particles.

In the design of accelerators in the past two decades, perhaps the most striking improvements have been made possible by the clever design of magnets so as to focus the circulating beam of particles more precisely on the predetermined orbit. The trick is to alternate around the orbit the gradients of the magnetic fields sensed by the circulating particles in two directions orthogonal to the orbit. The benefits of alternating gradient focusing are not merely the improvement of the quality of the beam but the reduction of the volume of the vacuum chamber which encloses the beam, the reduction of the gap between the poles of the magnets which follows and the simplicity that is as a result effected. One of the most striking visual features of the newer accelerators is the relative unobtrusiveness of their magnets.

Inevitably, the now demonstrated feasibility of superconducting magnets excites the interest of the machine builders, notably at Stanford University (separately from the electron accelerator SLAC) and at other places where particles are accelerated. One possibility is that the ultimate particle energy from existing accelerators might be increased by the replacement of electromagnets with superconducting magnets, while machines such as LEP, on which work has begun at CERN in Geneva, have been designed from the outset to accommodate superconducting magnets at some stage. The benefits of success, at least with closed-orbit machines, is that the civil engineering investment in a tunnel can be used to generate more energetic particles. □

Magnets for detection

THE use of magnets in producing beams of high-energy particles is obviously indispensable, but is similarly essential for their detection. Cosmic-ray physicists have long known this, and magnetic fields have been used as momentum spectrometers since cloud chambers came into common use.

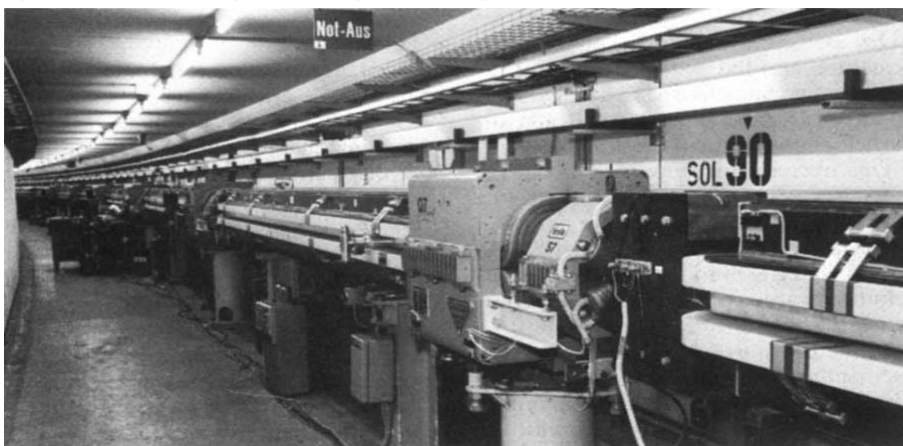
The principle is that of the mass spectrometer except that it is rarely necessary to arrange that all particles of the same mass, whatever their velocity, should be brought to a common focus. It is sufficient that the strength of the magnetic field be known, and the momentum inferred from the observed curvature of a track. The first V-particles, now with hindsight recognized as the first mesons containing a strange quark, were found in 1950 by installing a sensitive cloud chamber and a powerful electromagnet on the Pic du Midi in the Swiss Alps.

Nowadays, the same principle that the deflection of a charged particle in a magnetic field is inversely proportional to its momentum is usually a necessary part of particle detectors. Given that, in spark or bubble chambers for example, particles are "observed" individually, a measurement of momentum provides as much information as could be asked for. But the technology of spanning bubble chambers up to 2 metres in dimension with a more or less uniform magnetic field as strong as possible, remains a challenge.

The intricacies of this technology are perhaps most evident in the "beam switchyard" at the business end of the Stanford linear electron accelerator. Analogous to a shunting yard on a railway system, the hangar in which the system is contained is populated by a variety of electromagnets intended to channel successive bunches of electrons into predetermined vacuum lines.

The main problem the magnet designers face is that of arranging that the insulation between laminations will not be broken down by the accumulated dose of radiation. One common solution is that of coating a glass tape with epoxy-resin loaded with alumina, but the technology of protecting lamellar insulation from radiation damage is in its infancy.

The problem may be economically more significant if working thermonuclear plants ever come to depend more than incidentally on magnetic systems. The most challenging systems are the magnetic mirror machines in which charged particles constituting a hot plasma bounce repeatedly from one end to the other. But here, as in the manipulation of hot plasma in what is otherwise a vacuum, what will matter most is the configuration of the magnetic field.



The magnet line at Petra, the electron synchrotron at DESY, Hamburg.